AFRL-RY-WP-TR-2009-1331

MITIGATION APPROACHES FOR OPTICAL IMAGING THROUGH CLOUDS AND FOG

Mohsen Kavehrad

The Pennsylvania State University

NOVEMBER 2009
Final Report

Approved for public release; distribution unlimited.

See additional restrictions described on inside pages

STINFO COPY

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

*//Signature//   //Signature//

BRIAN K. STADLER   ROBERT J. FELDMANN, Chief
Project Engineer   EO Combat ID Technology Branch
EO Combat ID Technology Branch   EO Sensor Technology Division
EO Sensor Technology Division

//Signature//

COL. BRIAN C. FORD
Chief, EO Sensor Technology Division
Sensors Directorate

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YY)* November 2009	2. REPORT TYPE Final	3. DATES COVERED *(From - To)* 14 July 2008 – 14 November 2009

4. TITLE AND SUBTITLE MITIGATION APPROACHES FOR OPTICAL IMAGING THROUGH CLOUDS AND FOG	5a. CONTRACT NUMBER
	5b. GRANT NUMBER FA8650-08-C-7850
	5c. PROGRAM ELEMENT NUMBER 69199F
6. AUTHOR(S) Mohsen Kavehrad	5d. PROJECT NUMBER ARPR
	5e. TASK NUMBER YJ
	5f. WORK UNIT NUMBER ARPRYJ0B

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Pennsylvania State University Center for Information and Communication Technology Research (CICTR) 229 Electrical Engineering West University Park, PA 16802-2705	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSORING/MONITORING AGENCY ACRONYM(S)
Air Force Research Laboratory Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force	Defense Advanced Research Projects Agency/ Strategic Technology Office (DARPA/STO) 3701 N. Fairfax Drive Arlington, VA 22203-1714	AFRL/RYJM
		11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RY-WP-TR-2009-1331

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: 88ABW-09-5069; Clearance Date: 07 Dec 2009. This report contains color.

14. ABSTRACT

This report offers novel designs for active optical imaging through clouds and fog, motivated by many military applications, such as laser range-finder, target designation for seekers and laser radar.

The report focuses on jointly-optimized wireless optical multi-input multi-output (MIMO) transceivers for through-clouds airborne active optical imaging systems. Photolithographic techniques are employed in designing the optical transceiver subsystems. Our design goal is to select the diversity receiver branches field-of-view (FOV) in a way that the effect of multi-scattering dispersion is reduced by as much as possible. The light beam is converted to multiple beams for transmission over several independent wireless optical channels. At the optical imaging receiver platform, combined received beams are photo-detected, and an electronic receiver is employed to recover the images. We demonstrate high-resolution imaging can be achieved with a high reliability everywhere within the coverage area.

15. SUBJECT TERMS
MIMO

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT: SAR	18. NUMBER OF PAGES 194	19a. NAME OF RESPONSIBLE PERSON (Monitor) Brian K. Stadler
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER *(Include Area Code)* N/A

i

Table of Contents

List of Figures

List of Tables

Executive Summary

Under AFRL contract FA8650-08-C-7850, this report offers novel designs for active optical imaging through clouds and fog, motivated by many military applications, such as laser range-finder, target designation for seekers and laser radar.

The report focuses on jointly-optimized wireless optical multi-input multi-output (MIMO) transceivers for through-clouds airborne active optical imaging systems. Photolithographic techniques are employed in designing the optical transceiver subsystems. Our design goal is to select the diversity receiver branches Field-of-View (FOV) in a way that the effect of multi-scattering dispersion is reduced by as much as possible. The light beam is converted to multiple beams for transmission over several independent wireless optical channels. At the optical imaging receiver platform, combined received beams are photo-detected, and an electronic receiver is employed to recover the images. We demonstrate high-resolution imaging can be achieved with a high reliability everywhere within the coverage area.

In airborne active optical imaging systems, characterization of light propagation through clouds is essential to understanding and designing appropriate systems. We at Penn State have developed a Monte-Carlo-Ray-Tracing (MCRT) computer simulation modeling tool to investigate light propagation through different cloud obscurant media. We have used this tool plus extension tools we have developed, based on Markov modeling, to assess our proposed mitigation techniques for this application.

From channel characterizations made for RF Multi-Input-Multi-Output (MIMO) systems, we know that characteristics are heavily feature-dependent. This is due to the multiple images produced from different angles by having multiple beams departing the transmitter. These are provided to direction diversity receivers via multiple spatial transmissions.

We originated and designed concepts for a *Multi-Input-Multi-Output (MIMO)* wireless optical architecture referred to as *Multi-Spot Diffuse (MSD) configuration* with Multi-element optical transmitters and multi-branch optical receivers, proposed in [7]. We intend to use such configurations in this research. Professor Joseph Kahn of Stanford University, in a nice overview paper [73] describes our contributions in [7], as follows:

- Implementation of multi-branch angle diversity using non-imaging elements requires a separate optical concentrator for each receiving element, which may be excessively bulky and costly. Yun and Kavehrad proposed the fly-eye receiver [7], which consists of a single imaging optical concentrator (e.g., a lens) that forms an image of the received light on a collection of photo-detectors, thereby separating signals that arrive from different directions. In this article, we refer to this design as an imaging angle-diversity receiver, or simply an imaging receiver. Implementation of an angle-diversity receiver using imaging optics offers two advantages over a non-imaging implementation. First, all photo-detectors share a common concentrator, reducing size and cost. Second, all the photo-detectors can be laid out in a single planar array, facilitating the use of a large number of receiving elements or pixels.

- In non-Line-of-Sight (LoS) wireless optical links, Yun and Kavehrad [7] also proposed the spot-diffusing transmitter, which utilizes multiple narrow beams pointed in different directions, as a replacement for the conventional diffuse transmitter, which utilizes a single broad beam aimed at an extended reflecting surface. In this article, we refer to the spot-diffusing transmitter as a multi-beam or quasi-diffuse transmitter. While the diffuse transmitter provides considerable

immunity against beam blockage near the receiver, it yields a high path loss. The quasi-diffuse transmitter is expected to reduce path loss compared to the diffuse transmitter, because the narrow beams experience little path loss traveling from the transmitter to the illuminated reflective surfaces.

Later in [54], Kavehrad and Jivkova explained in detail the designs and implementation aspects of the two concepts described above.

Narrow wireless optical beams can minimize interference; subsequently making these links difficult to tap, thus providing inherent security which has been a point of major concern and a focus point for current research efforts. However, wireless optical links are most vulnerable to atmospheric particles, and relatively less vulnerable to rain. A great challenge in providing a near continuous availability for optical wireless links is their vulnerable nature to obstruction by fog, cloud, and dust particles. We note that, pure atmosphere represents a relatively clear medium for a transiting beam of light, typically 1dB of loss per kilometer (horizontally), is the accepted value. In vacuum, this transmission distance can grow to millions of kilometers and is really only limited by the inverse square law of spherical wave power dispersion. Photons tend to travel in a straight line until they hit something, are absorbed, reflected or refracted. Water as liquid droplets (rain, fog and snow) suspended in the atmosphere is a medium that presents three loss mechanisms in which; two are optical, and one is molecular. These are classified as geometric scattering, Mie scattering and molecular absorption. Liquid water attenuation for 1.5-micron light is very many decibels per meter. The common thread among these attenuation mechanisms is the atomic/molecular coupling properties between light and matter - more specifically, the light-waves electrical field and the molecules dipole (in the case of water).

Current research in wireless communication has focused on spatial-temporal diversity techniques (including coding) in an effort to increase transmission capacity and quality-of-service. Traditional modes of diversity include time, frequency, polarization, space, angle, field and multi-path. In digital communications applications [2], we demonstrated diversity through spectral encoding applied to ultra-short laser light pulses, using wavelet pulse shaping, in order to increase the robustness against temporal intersymbol-interference over the highly dynamic optical cloudy channels. We have established that using wavelet-shaped ultra-short pulses through clouds (multi-scattering channels) [2], it is possible to have transmission with a higher optical thickness compared to transmission of conventional non-return-to-zero square pulses, for the same target performance and transmission range. In other words, wavelet shaped ultra-short pulse streams propagate through thin clouds with less degradation than conventional pulse streams at the same transmission speeds. Using Fractal modulation, each receiver has a menu to choose the best received signal transmission rate, thus adaptation is feed-forward. At the same time, wavelets have the desirable properties of being both time and frequency limited, thus are able to pack a large amount of power in very short pulses, in addition to providing inherent diversity. A 100 fs pulse at 100 mJ would produce a peak power of 1 Terawatt.

Matched filtering can then be applied through Spatial Encoding and Decoding of frequency components of the broadband ultra-short light pulses. The encoding/decoding may be realized all-optically through photolithographic masks or other types of spatial light modulators. Some of these techniques might be suitable to the current investigation, namely imaging through clouds and fog, as they can be utilized against spatial intersymbol-interference. We will investigate these issues in this report.

A convenient way to produce ultra-short pulses of laser light is to employ mode-locked lasers. Many applications require ultra-fast optical waveform generation according to user specifications. This can be accomplished through manipulations on amplitude and phase of frequency elements in an ultra-short laser pulse spectrum. Optical pulse shaping techniques exploit the connection between frequency content and temporal shape of a signal. Basically, a waveform is synthesized by parallel modulation (or encoding) of spatial optical wavelength components that make up the ultra-short pulse.

We originally intended to apply this approach on the imaging system. However, if beam splitting is to be done on the shaped ultrashort pulses, it would distort these pulses, as splitting the light beam through a photolithographic beam splitter acts as linear filtering and band limiting the ultrashort pulses which happen to be extremely broadband. So, instead, the splitting is done prior to mode-locking, on the pump laser which is a CW laser.

The space-time focusing properties of this approach leads to a new class of *"Opportunistic"* imaging systems with significant advantages over current RF approaches. The spatial focusing potential of this approach is an appealing quality, in power saving. Interference issues of shared RF bands are non-existent here.

Innovative Claims

This report is focused on a novel design of a jointly-optimized wireless optical MIMO transceiver system for broadband through-cloud multi-channel optical imaging. Photolithographic techniques are employed in designing the optical transceiver subsystems. The source light is converted to multiple beams for transmission over several independent wireless optical channels. At the receiver platform, combined received beams are photo-detected, and an electronic receiver is employed to recover the images. A high resolution and a high reliability everywhere within the coverage area can be expected. To achieve this, we worked on novel subsystems designs as listed below:

- *Computer simulation modeling of the cloud channel;*
- *Mathematical analysis of the MIMO imaging performance over the coverage area;*
- *Computer simulation evaluation of the MIMO imaging system resolution and performance;*
- *Investigation of the effects of turbulence and scintillation on imaging performance;*
- *Performance evaluation of adaptive optics compensation techniques on imaging resolution;*
- *Investigation of different types of deconvolution techniques that are able to reconstruct scintillation-degraded blurred images;*
- *Numerical link performance assessment with transmit power and transmission rate as variables of the computations;*
- *Identify the number of transmitted beams for satisfactory link performance;*
- *Identify the best number of receiver branches and the associated FOV of each branch;*
- *Several types of curved holographic optical mirrors can potentially be employed at the receiver. We compared the performance of spherical mirror on a spherical substrate, a spherical mirror on a flat substrate, a parabolic mirror on a flat substrate;*

- *Design and fabrication of a computer generated photolithographic beam-splitter, used at the transmitter for producing multiple collimated transmitted light beams within a controlled solid angle;*
- *Design and fabrication of a Holographic Optical Element (HOE) for combining the received light concentration and filtering functions, in order to create a reasonably wide FOV angle-diversity photo-receiver;*
- *Determined the relation between the solid-angle of the transmitted light beam and the channel diversity requirements, for a specified performance requirement;*
- *Examined the impact of different operational wavelengths on the hologram efficiency.*

In comparison with any currently available imaging system, this research and design effort has resulted in higher resolution images and a larger range of coverage. The imaging system configuration is tolerant to shadowing and blockage. Also, the optical medium is more secure than the RF counterparts against any intrusion.

Chapter 1
Introduction

1.1   Motivation

Since the invention of Laser in 1960, it has found multiple applications including communications, remote sensing, and imaging. The advantages of performing imaging in the optical band are manifold. Modern Lidar and Ladar systems are preferred over their RF counterparts due to the higher resolution, faster area search rate, and ease of human interpretation. All these benefits are direct results of the smaller wavelength of optical waves. Medical imaging benefits from the ability of laser to penetrate soft tissues and organs, for detecting tumors and cancerous glands. Active imaging using laser beams is also of particular importance to the military community. Laser imaging systems find use in ground battlefields as well as air-to-ground surveillance. The ability of obtaining high-powered short-pulsed laser modules today has transformed our abilities to image through aerosols and turbulence. One good example of the use of laser for air-to-ground surveillance is the Synthetic Aperture Ladar for Tactical Imaging (SALTI) [1], which is an airborne interferometric synthetic advanced laser radar imager capable of producing high-resolution three-dimensional imagery at long ranges. The advantage of active imaging is particularly noted in its daytime and nighttime use, control over the pattern and strength of the emanating laser from the transmitter, and incorporation of parallel schemes.

The challenges associated with active laser imaging are also manifold. The success of the SALTI was in clear atmosphere, but the same approach does not necessarily apply to clouds and turbulent atmosphere, which is a more likely scenario in long-distance imaging case. In contrast to radio wavelengths, the optical wavelengths are of the same order of particles size as vapor; therefore, atmospheric phenomena such as cloud, fog and haze interact with laser beams and scatter them [2]. As a result of multiple scattering events, the transmitted laser beam is broadened and the energy is redistributed in time and space. This atmospheric attenuation imposes a big challenge on laser imaging systems, and it can be as severe as 300 dB/km in heavy fog [3]. As a result, the light reaching the target to be imaged undergoes attenuation, and similar attenuation is also present on the return path. Furthermore, in an optical imaging system, back-scattered light from the scattering media impinges on the receiver and gives rise to a steady background noise and limits contrast.

Multiple scattering is not the only degrading factor. Temperature and wind speed variations in different layers of the atmosphere contribute to turbulence, which introduces refractive index fluctuations along the optical path. This causes amplitude and phase fluctuations in the received wave-front [4][5][6]. Optical imaging systems experience severe resolution degradation due to PSF broadening [6]. Fig. 1.1 shows the issues that an optical imaging system should resolve.

<div align="center">

(a)   **(b)**

</div>
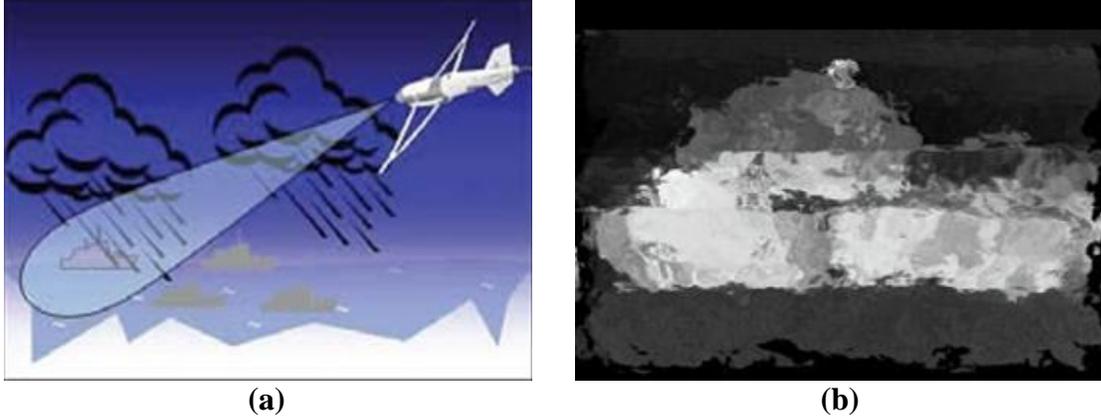
Fig. 1.1: (a) Backscattering and (b) PSF broadening in an Optical imaging system.

1.2   Objective

To exploit the great potential of optical imaging systems, a comprehensive knowledge of the channel is required. Furthermore, various simulation tools must be developed to model the optical channel and imitate the performance limiting factors. Given such a channel model, special measures should be taken in design of any optical imaging and communications system to make it sustainable to possible sources of degradation. Such measures are necessary in both hardware and software design. In other words, both transmitter and receiver physical configurations should be chosen to optimize system performance. Additionally, corrective modules and proper post-processing blocks should be considered in the system design.

Multiple-Input Multiple-Output (MIMO) configuration is proved to be effective in fading RF channels. Moreover, multi-spot diffuse configuration, first introduced by Yun and Kavehrad [7], has a great potential to be exploited in optical imaging, as discovered by MIT Lincoln Lab [8][9][10]. In this report, we propose MIMO imaging systems and investigate their performance under various atmospheric conditions. Then, using restoration techniques, we try to find a solution for performance degrading atmospheric phenomena.

1.3   Organization

First and foremost, we investigate the mechanisms of scattering and absorption and study theories that describe these phenomena in chapter 2. Next and in chapter 3, based on Mie theory, Monte Carlo Ray tracing (MCRT) algorithm is implemented and some of the channel parameters are found through post-processing of the simulation results. In chapter 4, by direct extraction of classical Markov chain, associated with angular distribution, evolution of laser beam and Second Largest Eigen Module (SLEM) of Monte Carlo Markov Chain, feasibility and achievable rates in FSO are investigated. Next in chapter 5, atmospheric turbulence models are studied and different methods of generating phase screens are reviewed and compared based on accuracy and computational complexity. Chapter 6 discusses MIMO wireless optical communications and illustrates Bit Error Rate (BER) performance improvements obtained in an Intensity Modulation/ Direct Detection (IM/DD) optical channel. Furthermore, a Single-Input Single-Output (SISO) configuration equipped with Adaptive Optics (AO) correction module is simulated and compared to a MIMO equivalent system. Chapter 7 introduces a spatially-multiplexed MIMO imaging system, inspired by multi-spot diffuse indoor wireless optical communication scheme. Given the comprehensive channel model developed in chapters 2 to 5, performance limitations and challenges that such

an imaging system faces in an atmospheric channel are explored. Chapter 8 demonstrates performance improvements achieved using a combination of image restoration techniques, such as time-gating, Blind Deconvolution (BD) and AO. Finally chapter 9 concludes this document and outlines possible future work.

Chapter 2
Mie Scattering Theory

2.1   Introduction

Light propagation through clouds is basically a multiple-scattering problem in which beams or photons undergo several scatterings before either exiting the cloud or reaching the receiver. Most of the particles in scattering media such as atmosphere and ocean have dimensions comparable to (or less than) optical wavelengths. As a result, the study of optical signal propagation in such media needs some knowledge of Mie scattering theory. This theory is "the application of the Maxwell's equation to the problem of a homogeneous sphere radiated by a plane wave from a single direction" [11].

2.2   Single Scattering

In calculation of scattering from aerosols, it is usually assumed that the particles are spherical. This assumption is reasonable as a result of random orientation of particles. Nevertheless, some authors have addressed the problem of Mie scattering from non-spherical particles [12]. Fig. 2.1 illustrates the underlying geometry of Mie theory [13].

Assume a spherical particle of diameter D with refractive index m_i is suspended in a medium with refractive index m_o. The relative refractive index is $m=m_i/m_o$ and the wavelength of the incident beam is λ. The incident electric field can be expressed as

$$E_i = E_0 e^{-jkz+j\omega}$$

(2.1)

where E_0 is amplitude of the incident field, and $k = \dfrac{2\pi}{\lambda}$ is the wave number and the wave is traveling in $+\vec{z}$ direction. The incident electric field can be decomposed to a parallel and a perpendicular component to the scattering plane [14].

(2.2)
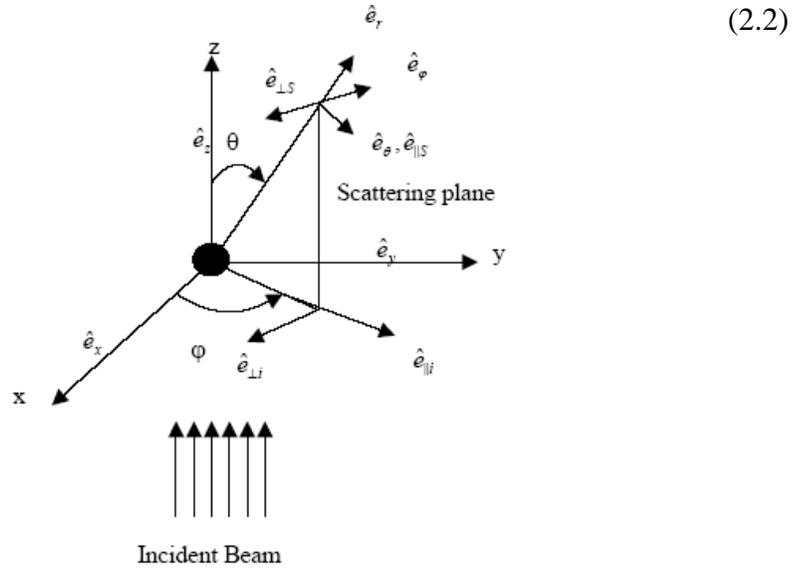
Fig. 2.1: Illustration of geometry for calculation of Mie scattering.

The solution for the scattered electric field is given in terms of two scalar components $S_1(\theta)$ and $S_2(\theta)$, which are functions of perpendicular and parallel-polarized components of incident field with respect to scattering plane (the plane in which θ is measured) [13][14][15].

$$S_1(\theta) = \sum_{n=0}^{\infty} \frac{2n+1}{n(n+1)} (a_n \pi_n + b_n \tau_n) \tag{2.3}$$

$$S_2(\theta) = \sum_{n=0}^{\infty} \frac{2n+1}{n(n+1)} (b_n \pi_n + a_n \tau_n) \tag{2.4}$$

where

$$a_n = \frac{\psi_n(x)\psi_n'(mx) - m\psi_n'(x)\psi_n(mx)}{\xi_n(x)\psi_n'(mx) - m\xi_n'(x)\psi_n(mx)} \tag{2.5}$$

$$b_n = \frac{m\psi_n(x)\psi_n'(mx) - \psi_n'(x)\psi_n(mx)}{m\xi_n(x)\psi_n'(mx) - \xi_n'(x)\psi_n(mx)} \tag{2.6}$$

and

$$\psi_n(x) = xj_n(x) = \sqrt{\frac{\pi x}{2}} J_{n+\frac{1}{2}}(x) \tag{2.7}$$

$$\xi_n(x) = xh_n(x) = \sqrt{\frac{\pi x}{2}} (J_{n+\frac{1}{2}}(x) + iY_{n+\frac{1}{2}}(x)) \tag{2.8}$$

Note that, $J_n(x)$ and $Y_n(x)$ are the ordinary Bessel functions of first and second kind, respectively. Also, $j_n(x)$ is the spherical Bessel function of first kind while $h_n(x)$ is the Spherical Hankel function of first kind. Moreover, the argument x in the above equations is the size parameter and is equal to:

$$x = \frac{2\pi n}{\lambda} \tag{2.9}$$

and n is the refractive index of the particle; Furthermore, τ_n and π_n in (2.3) and (2.4) are as follows:

$$\pi_n = \frac{P_n^1(\cos\theta)}{\sin\theta} \tag{2.10}$$

$$\tau_n = \frac{\partial P_n^1(\cos\theta)}{\partial\theta} \tag{2.11}$$

where $P_n^1(\cos\theta)$ is the first order associated Legendre function. Also, we have the following recursive equation for π_n and τ_n.

$$\pi_n = \frac{n-1}{2n-1}\cos\theta\pi_{n-1} - \frac{n}{n-1}\pi_{n-2} \tag{2.12}$$

$$\tau_n = n\cos\theta\pi_n - (n+1)\pi_{n-1} \tag{2.13}$$

Given $S_1(\theta)$ and $S_2(\theta)$ from (2.3) to (2.13), we can finally compute the scattered electric field as:

$$\begin{bmatrix} E_{\|S} \\ E_{\perp S} \end{bmatrix} = \frac{e^{jk(R-z)}}{-jkR} \begin{bmatrix} S_2(\theta) & 0 \\ 0 & S_1(\theta) \end{bmatrix} \begin{bmatrix} E_{\|i} \\ E_{\perp i} \end{bmatrix} \tag{2.14}$$

From these equations, one can see that basis vectors are different for incident and scattered fields (see Fig. 2.1). Furthermore, since the transfer matrix is diagonal, the incident field does not depolarize. The effect of polarization is better understood by studying the *Stokes* vector *[I,Q,U,V]*, the elements of which are given by:

$$I = <|E_\||^2 + |E_\perp|^2>$$

$$Q = <|E_\||^2 - |E_\perp|^2>$$

$$U = <E_\| E_\perp^* + E_\perp E_\|^*> \tag{2.15}$$

$$V = <E_\| E_\perp^* - E_\perp E_\|^*>$$

With this notation, the relationship between the Stokes vectors of the incident and scattered fields becomes

$$\begin{bmatrix} I_S \\ Q_S \\ U_S \\ V_S \end{bmatrix} = \frac{1}{k^2 R^2} \begin{bmatrix} S_{11} & S_{12} & 0 & 0 \\ S_{12} & S_{22} & 0 & 0 \\ 0 & 0 & S_{33} & S_{34} \\ 0 & 0 & -S_{34} & S_{33} \end{bmatrix} \begin{bmatrix} I_i \\ Q_i \\ U_i \\ V_i \end{bmatrix} \tag{2.16}$$

where the (Mueller) matrix elements are defined as:

$$S_{11} = \frac{1}{2}(|S_2|^2 + |S_1|^2)$$

$$S_{12} = \frac{1}{2}(|S_2|^2 - |S_1|^2)$$

$$S_{33} = \frac{1}{2}(S_2^* S_1 + S_2 S_1^*) \tag{2.17}$$

$$S_{34} = \frac{1}{2}i(S_1 S_2^* - S_2 S_1^*)$$

Stoke formulation is useful since it enables us to apply superposition principle to the Stokes parameters of the light scattered by a collection of randomly separated particles. From the above formulation, it can be seen that an incident wave with parallel polarization results in a scattered radiation with parallel polarization. This is the same for the case of perpendicular

polarization. The energy removed from the incident beam by scattering on the particle is proportional to the extinction cross section C_{ext}.

$$C_{ext} = C_{abs} + C_{sca} \qquad (2.18)$$

which is the sum of the absorption and scattering cross-sections. The extinction cross section is

$$C_{ext}(r) = \frac{4\pi}{k^2} \text{Re}\{S_i(0,r)\}, \qquad (2.19)$$

where $i = 1$ or 2. (Note: $S_1(0,r) = S_2(0,r)$). The scattering cross-section is expressed as:

$$C_{sca}(r) = \frac{1}{k^2} \int_{4\pi} S_{11}(\theta,r) d\omega \qquad (2.20)$$

and the integration is over the entire solid angle. The single scattering albedo α (or in some texts ϖ) is the ratio of scattering and extinction cross-sections.

$$\alpha = \frac{C_{sca}}{C_{ext}} \qquad (2.21)$$

Finally, backscattering cross section is given by

$$Q_{back} = \frac{4\pi}{k^2} S_{11}(180^0) \qquad (2.22)$$

By normalizing the above scattering and extinction cross-sections to the geometrical cross-sections, one may define the corresponding efficiencies.

$$Q_{sca} = \frac{C_{sca}}{\pi r^2} \qquad (2.23)$$

$$Q_{ext} = \frac{C_{ext}}{\pi r^2} \qquad (2.24)$$

$$Q_{back} = \frac{C_{back}}{\pi r^2} \qquad (2.25)$$

The scattering direction is defined via a three-dimensional probability density function known as the *Phase Function* $P(\theta)$. The phase function is given by:

$$P(\theta) = \frac{4\pi S_{11}(\theta,r)}{k^2 C_{sca}(r)}, \qquad (2.26)$$

where it is normalized so that its integral over the entire solid angle is 4π [16], i.e.:

$$\int_{over 4\pi} P(\theta)d\Omega = \int_0^\pi \int_0^{2\pi} P(\theta) \sin(\theta) d\theta d\varphi = 4\pi \qquad (2.27)$$

Note that, $S_{11}(\theta, r)$ encompasses both parallel and perpendicular polarizations as defined in (2.17). Hence, the above-mentioned phase function accounts for both parallel and perpendicular incident wave scattering. Fig. 2.2 and Fig. 2.3 demonstrate the phase function for the case $\dfrac{r}{\lambda} = 1$, while Fig. 2.4 and Fig. Fig. 2.5 show the phase function for $\dfrac{r}{\lambda} = 0.1$. It is clear from these figures that as the ratio of particle radius to wavelength increases, the scattering becomes more anisotropic.
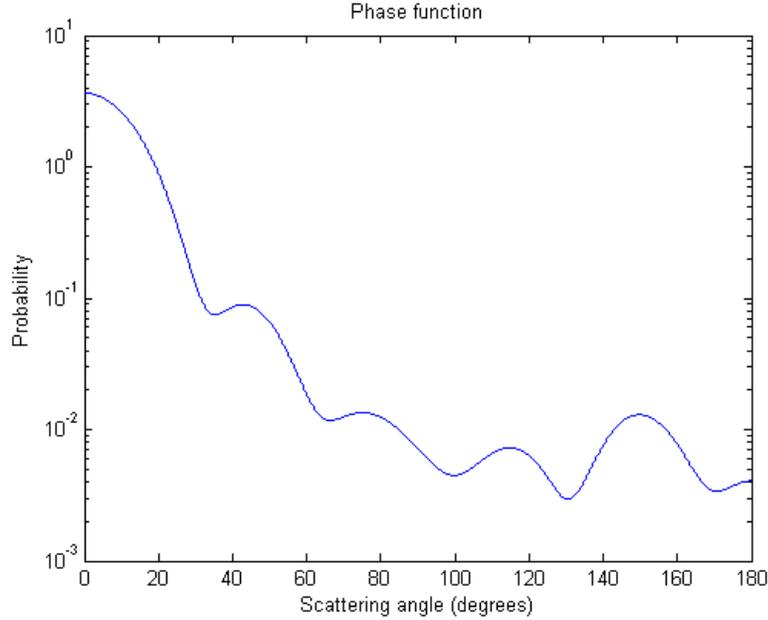
Fig. 2.2: Phase function for $\dfrac{r}{\lambda} = 1$

2.3   Poly-dispersion

Clouds are composed of many particles from various sizes. Thus, we have to generalize the above concepts and take into account the particle size distribution of a scattering medium. To this end, we define a scattering cross-section per unit volume β_{sca}, expressed as:

$$\beta_{sca} = \int_0^\infty C_{sca}(r)n(r)dr \qquad (2.28)$$

where the function *n(r)* is the particle size distribution function defined as the number of particles per unit volume with radii between *r* and *r+dr*.
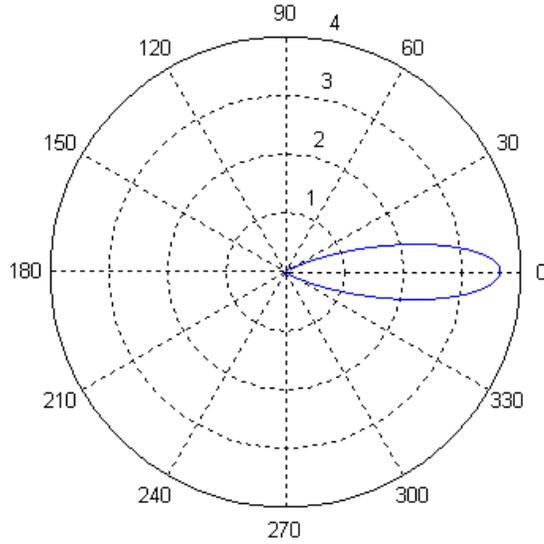
Fig. 2.3: Phase function for $\dfrac{r}{\lambda} = 1$ (Polar Coordinate).

These distribution functions are usually expressed in units of cm^{-3} μm^{-1}. Also, note that the scattering cross-section per unit volume has the units of inverse length. It is often called the scattering "coefficient". The extinction coefficient is defined similarly and the corresponding single-scattering albedo is defined as $\alpha = \beta_{sca} / \beta_{ext}$. The corresponding generalization of the Mueller matrix elements are:

$$S_{ij}(\theta) = \int_{0}^{\infty} S_{ij}(\theta, r) n(r) dr .$$

(2.29)

Note that, the elements are now defined per unit scattering volume. Also, we can generalize the phase function to the case of poly dispersion as follows:

$$P(\theta) = \frac{4\pi S_{11}(\theta)}{k^2 \beta_{sca}}$$

(2.30)

13
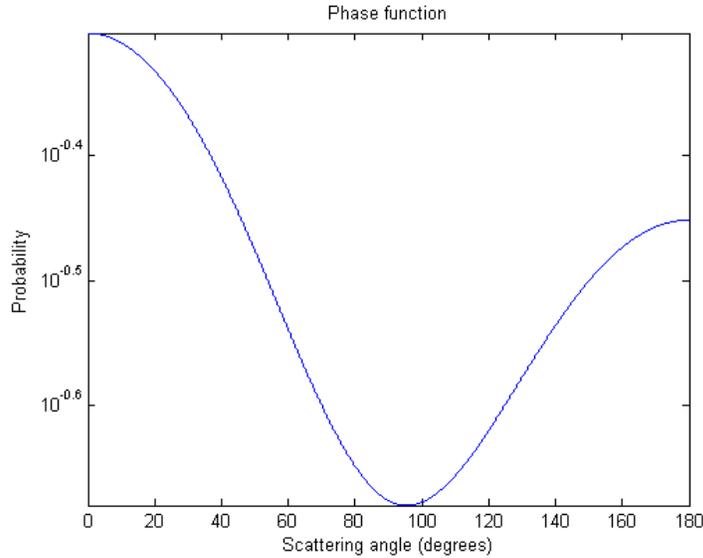
Fig. 2.4: Phase function for $\dfrac{r}{\lambda} = 0.1$.
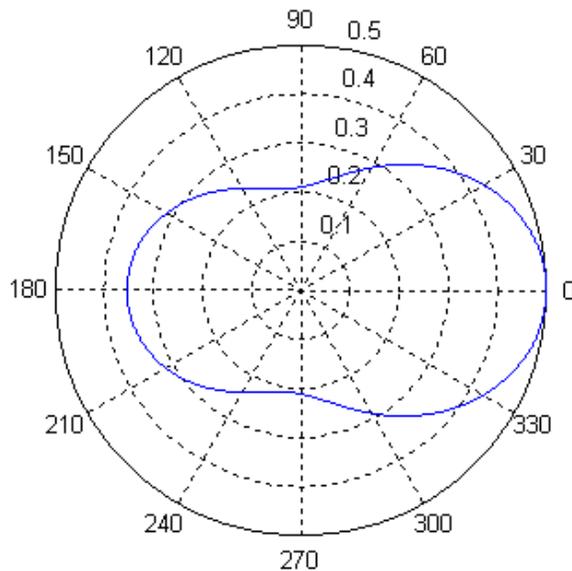
Fig. 2.5: Phase function for $\dfrac{r}{\lambda} = 0.1$ (Polar Coordinate).

2.4 Particle-size Distribution Models

Clouds and aerosols are frequently a combination of more than one type; each type may have its own relative density distribution and refractive index properties. Each particle has a particulate bulk density, ρ, in units of g cm^{-3}, or as a dimensionless specific gravity. Moreover, each species may be characterized by a given mass concentration, C (g cm^{-3}), which represents the weight of lofted material mixed within a unit volume of air. The quantities C and ρ are related by particle size distribution, denoted as $n(r)$. The total number

14

of particles per unit volume, Number Density, is denoted by N, which is related to particle size distribution through:

$$N = \int_0^\infty n(r)dr \,, \tag{2.31}$$

where, as mentioned before, $n(r)$ has the units of particle per cm^3 μm. We may now write mass concentration C, as a function of $n(r)$ and ρ, as:

$$C = \frac{4}{3}\pi \int_0^\infty n(r)r^3\rho dr \,, \tag{2.32}$$

where $\frac{4}{3}\pi r^3 \rho$ is the mass of particle of radius r . There are a number of analytic models of aerosol size distributions in current use. These include gamma, modified gamma, and lognormal.

2.4.1   Gamma Distribution

Gamma Distribution, used by snow and rain models, is given by:

$$n(r) = N_0 r^m \exp(-\Lambda r) \,, \tag{2.33}$$

where N_0 is the y intercept, Λ is the slope, and m is the order of gamma size distribution. Taking $m = 0$, we find the number density to be

$$N = N_0 / \Lambda \,. \tag{2.34}$$

The mass concentration for this distribution can be expressed as:

$$C = 4\pi N_0 \rho \Gamma(4)/(3\Lambda^4) = 8\pi N_0 \rho / \Lambda^4 \,. \tag{2.35}$$

2.4.2   Modified Gamma Distribution

Modified Gamma (MG) Distribution, first introduced by Deirmendjian [17] and used by fog models, is given by:

$$n(r) = ar^\alpha \exp(-br^\gamma) \,. \tag{2.36}$$

Defining the mode or modal radius as $r_c^\gamma = \dfrac{\alpha}{b\gamma}$, we can rewrite (2.36) as:

$$n(r) = r_c r^\alpha \exp(-(\alpha/\gamma)(r/r_c)^\gamma) \,. \tag{2.37}$$

In this case, the total number density is

$$N = a\gamma^{-1}b^{-(\alpha+1)/\gamma}\Gamma\left(\frac{\alpha+1}{\gamma}\right) \,. \tag{2.38}$$

The mass concentration for this distribution can be expressed as

$$C = \frac{4}{3}\pi\rho r_c^3 N \frac{\Gamma((\alpha+4)/\gamma)}{(\frac{\alpha}{\gamma})^{\frac{3}{\gamma}}\Gamma((\alpha+1)/\Gamma)} \, . \tag{2.39}$$

Table 2.1 shows the parameters of "Modified Gamma" Distribution for different kinds of cloud.

Table 2.1: Parameters of "Modified Gamma" distribution for various kinds of clouds.

Distribution Type	N (cm^{-3})	a	r_c	α	γ	b
Cumulus	250	2.604	6	3	1	0.5
Stratus	250	27.0	3.3333	2	1	0.6
Stratus/ strato-cumulus	250	52.734	2.6667	2	1	0.75
Stratocumulus	150	9.375	4	2	1	0.5
Alto-stratus	400	6.268	1.8	2	1	1.111
Nimbo-stratus	200	7.676	11.7647	5	1	0.425
Cirrus	0.025	2.21×10^{-12}	64	6	1	0.09375
Thin cirrus	0.5	0.011685	4	6	1	1.5

2.4.3　Lognormal Distribution

The lognormal distribution, used by the smoke, dust, rural, urban, maritime, tropospheric, desert, and the Navy aerosol model, is given by the equation (2.40):

$$n(r) = \sum_{i=1}^{k} \frac{N_i}{\sqrt{2\pi}\sigma_i r} \exp\left[-\frac{(\ln r - \ln r_{gi})^2}{2\sigma_i^2} \right], \tag{2.40}$$

where r_{gi} is the geometric mean radius (or mode radius) in μm, σ_i is the width of distribution, and N_i is the number density of the ith component. The mass concentration for this distribution is given by:

$$C_i = \frac{4}{3}\pi\rho N_i r_{gi}^3 \exp\left(9/2(\ln(10)\sigma_i)^2\right). \tag{2.41}$$

2.5　Phase Function

Phase function gives the directional distribution of radiation scattered by the particle. The phase function is the probability that incident wave is scattered through a scattering angle θ into an element of solid angle $d\Omega$. Given the above particle size distributions, we can derive the phase function for different kinds of clouds, using equations in sections 2.2 and 2.3. Fig. 2.6 demonstrates some of these phase functions. Since the azimuth scattering angle, φ, is uniformly distributed in $[0,2\pi]$, the *Phase Function* can be plotted only against the polar scattering angle, θ. The "normalized" phase function, as generally used in the literature, is normalized so that its integration over all possible scattering angles (*4π* steradians) is unity, thus qualifying it to be a PDF. With no φ dependence of $P(\theta)$, the integration over φ simply contributes a factor of *2π*. This leads to normalization definition for $P(\theta)$ of:

16

$$2\pi.\int_0^\pi \frac{P(\theta)}{4\pi}\sin(\theta)d\theta = \int_0^\pi P(\theta)\frac{\sin(\theta)}{2}d\theta = 1. \tag{2.42}$$

Thus, *PDF* of θ can be extracted from the 'normalized' phase function and be expressed as:

$$f(\theta) = P(\theta)\frac{\sin\theta}{2}. \tag{2.43}$$

Unfortunately, expressing the phase function versus θ caused a great deal of confusion in the published literature [18][19]. For example, in the case of isotropic scattering, by mistaking the "phase function" for *PDF* of θ, one may think that θ is uniformly distributed in *[0, π]*. However, it is the phase function (*PDF* of solid angle Ω) that is flat and θ is distributed as $f(\theta) = \frac{\sin\theta}{2}$ [18][19][20].
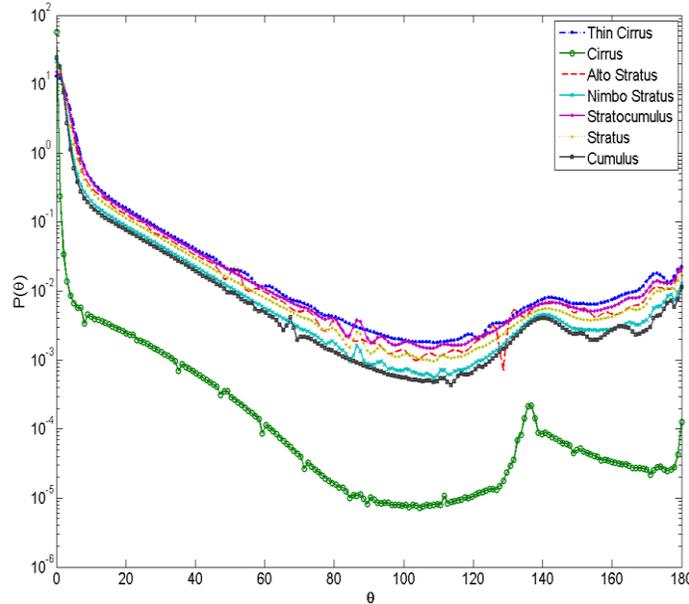
Fig. 2.6: Scattering Phase Function for different types of cloud.

When many scattering events have occurred and/or optical thickness is sufficiently large, the phase function can simply be approximated. One such approximation is the Henyey-Greenstein (HG) phase function [21],

$$P_{HG}(\theta, g) = \frac{1}{4\pi}\frac{1-g^2}{(1-2g\cos(\theta)+g^2)^{3/2}}, \tag{2.44}$$

where *g* is called the *asymmetry parameter* and is a measure of the ratio of forward to backward scattering. This phase function is interesting since it depends only on a single parameter *g*. For isotropic scattering *g* is *0* and for highly peaked scattering it approaches 1. Asymmetry parameter can also be negative in the case of back scattering. This parameter can be derived from (2.45).

$$g = \frac{\displaystyle\int_{4\pi} P(\theta)\cos(\theta)d\Omega}{\displaystyle\int_{4\pi} P(\theta)d\Omega} = \frac{\displaystyle\int_0^\pi\int_0^{2\pi} P(\theta)\cos(\theta)\sin(\theta)d\theta d\phi}{\displaystyle\int_0^\pi\int_0^{2\pi} P(\theta)\sin(\theta)d\theta d\phi}. \qquad (2.45)$$

In this project, we are mostly interested in cumulus clouds. While this cloud demonstrates a highly peaked phase function, which may be helpful from imaging perspective, it has the largest extinction coefficient as well, and can be considered as the "worst case" in this sense. For cumulus cloud g is 0.85. Fig. 2.7 compares the cumulus cloud phase function and the approximate phase function of g equal to 0.85. From this figure, one can see that although these phase functions have the same asymmetry parameter, g, they are completely different.
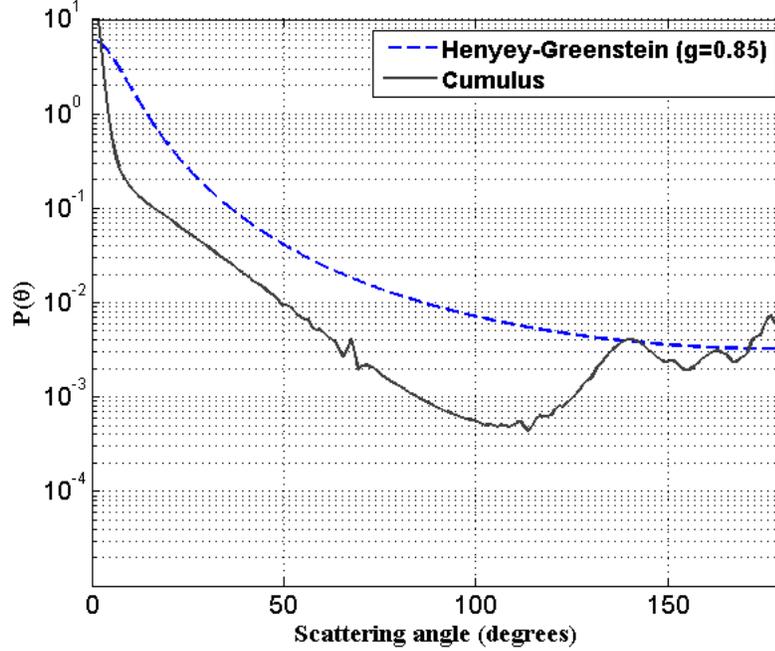
Fig. 2.7: Comparison of Cumulus cloud phase function with Henyey-Greenstein phase function of asymmetric parameter g=0.85.

One reason that causes these two different phase functions have the same asymmetry parameter is the definition of asymmetry parameter. As (2.45) shows, $P(\theta)$ is multiplied by $\sin(\theta)$ in both nominator and denominator. This causes the g parameter to be biased against highly peaked phase functions [23]. Also, note that the approximate phase function is monotonically decreasing and does not demonstrate back scattering phenomenon properly. As a result, since the exact function slightly increases at large angles to account for back scattering, its asymmetry parameter becomes the same as the approximate phase function, even though it has a stronger peak in forward direction. In this dissertation, we use exact phase function. It can be shown that the results using exact and approximate phase functions turn out to be quit similar. However, in some previous works, [24] and [18], it has been shown that approximate and exact phase functions generate different results in biomedical

applications. As will be illustrated later, this contradiction is a result of mistaking the phase function (or volume scattering function) with the PDF of θ [19].

2.6   Conclusions

In this chapter, single-scattering of light by particles is studied and parameters and functions involved in Mie theory are described. Mie Theory, although very useful in characterizing single scattering event, is not sufficient for describing multiple scatterings in a particulate medium. To explain wave propagation in a scattering medium, this theory should be combined with a powerful tool that can demonstrate consecutive scattering events. One such a tool is the Monte Carlo Ray Tracing (MCRT) algorithm, which basically implements a Markov Chain and tracks the path of a large body of photons in the propagation medium. Then, at a location of interest, i.e., receiver plane, statistics of impinging photons can be extracted via some post-processing. In the next chapter, we elaborate on this algorithm in details.

Chapter 3
Monte Carlo Ray Tracing through Cloud

3.1 Introduction

Airborne laser communication involves light propagation through atmospheric phenomena such as cloud, fog, haze, rain, or snow. In this process, transmitter launches laser pulses containing large number of photons. Due to the presence of particles in the atmosphere, laser pulse propagation through an optically thick medium is essentially a multiple-scattering problem, in which photons undergo several scattering before either escaping the medium or reaching the receiver. Since atmospheric conditions are always changing, we are dealing with a time-variant channel. This channel causes the optical signal to spread in space and time and produces *angular, spatial*, and *temporal* dispersions. To understand different aspects of this channel, both analytical methods and Monte Carlo simulations have been used. In this chapter, we study the multiple scattering problem and describe the Monte Carlo Ray Tracing algorithm as a means to explain light propagation through optically thick media.

3.2 Multiple Scattering

In chapter 2, single scattering from spherical particles is discussed using Mie theory. Moreover, different parameters of scattering such as phase function, extinction and scattering coefficients, and single scatter albedo are introduced. However, these parameters describe only a single scattering event. As mentioned earlier, light propagation through atmosphere, ocean, or any other optically thick medium is a multiple scattering problem. In other words, when photon is launched, it collides with several particles and is scattered or absorbed due to each collision.

Assume that photon is initially traveling in the *z* direction. As this photon interacts with an atmospheric particle, it is either absorbed with a probability of $1 - \varpi$ or scattered with a probability of ϖ, where ϖ is the single scatter albedo. If photon is scattered, the new traveling direction is governed by two random variables θ and ϕ, which are called polar and azimuthal angles of scattering, respectively. While ϕ has a uniform distribution in $[0, 2\pi]$, the distribution of θ is determined by the phase function, $P(\theta)$. The photon continues traveling in this new direction until it collides with another particle. Traveling distance between the two successive collisions is an exponential random variable with a mean equal to the inverse of scattering coefficient. In other words:

$$P(d) = \frac{1}{D_{ave}} \exp(-d / D_{ave}) \tag{3.1}$$

and $D_{ave} = \dfrac{1}{\beta_{sca}}$. When photon collides with the next particle, the new θ and ϕ are measured with respect to current traveling direction but not the initial *z* direction. Fig. 3.1 shows the geometry of multiple scatterings.

3.3 Monte Carlo Ray Tracing

Monte Carlo Ray Tracing (MCRT) tracks the path of a large number of photons from the transmitter to the receiver plane to get the statistical parameters required for channel

measurements. These parameters include the average cosine of angle of arrival or angular dispersion, the average distance from the centre of the receiver plane or spatial dispersion, and the maximum delay spread or temporal dispersion of the received light signal. MCRT continues tracking as many photons that the statistical fluctuations become negligible.

From Fig. 3.1 we can see that whenever a photon is scattered, its propagation direction changes. Accordingly, it can be assumed that the trajectory of a photon is a random walk in two dimensions. Since the new direction of propagation is defined relative to the previous traveling direction, photon propagation in such a medium can be described by a Markov chain, in which the previous traveling direction can be considered as the previous state and random variables θ and ϕ are the system input at each scattering event.
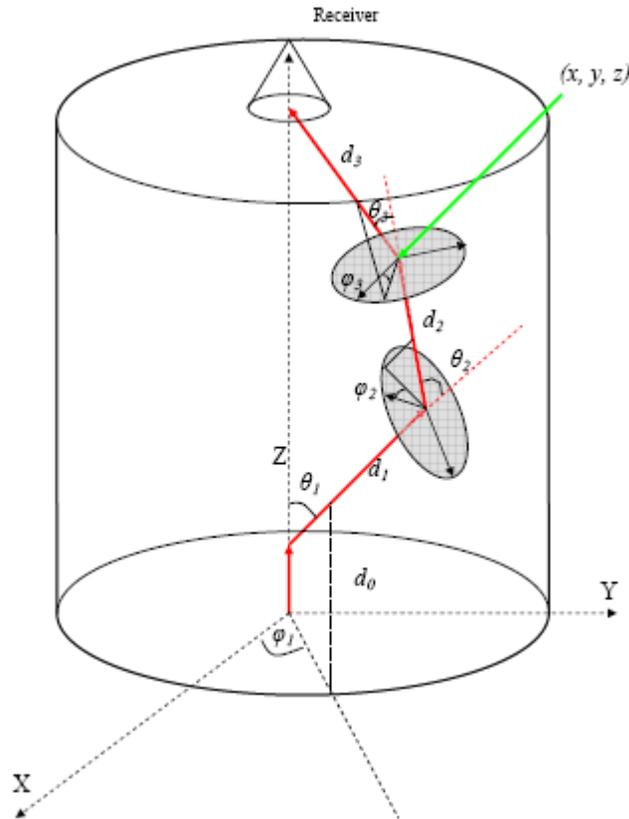
Fig. 3.1: Geometry of multiple scattering

As mentioned earlier, θ and ϕ are defined relative to previous traveling direction; therefore, we need a series of geometric transformations to figure out the position and traveling direction of a photon with respect to global coordinates after multiple scatterings. Since in each scattering event, first z and x axes are rotated with an angle θ, and then x and y axes are rotated with an angle ϕ, we have to combine the effect of these two rotations to derive the rotation matrix B for each scattering. First suppose that z and x axes are rotated with an angle θ, while y-axis is kept fixed. Fig. 3.2 shows this rotation. Any point in the new coordinate system (x', y', z') can be mapped onto (x, y, z) coordinate with the following geometric transformation [25].

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \Re_\theta \begin{bmatrix} x' \\ y \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y \\ z' \end{bmatrix} \qquad (3.2)$$

Now, suppose that axes x and y are rotated with an angle φ, while z-axis is kept fixed. Fig. 3.3 shows this rotation. One can relate any point in the new coordinate and the old one, as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \Re_\phi \begin{bmatrix} x' \\ y' \\ z \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z \end{bmatrix}. \qquad (3.3)$$
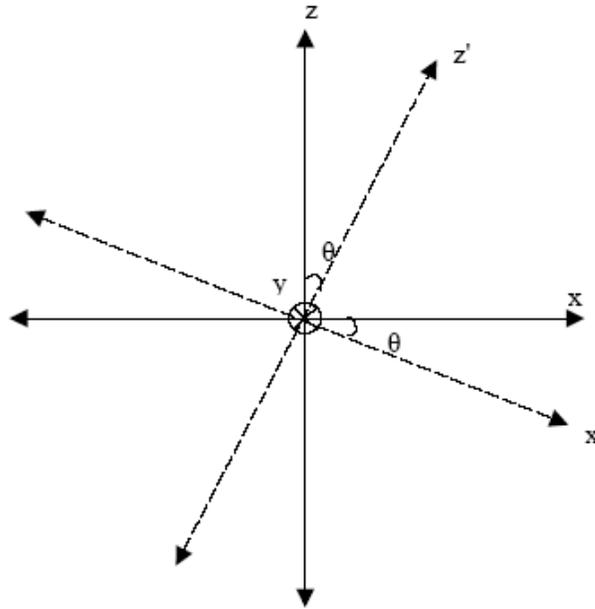
Fig. 3.2: Rotation axes z & x while keeping the y-axis fixed.
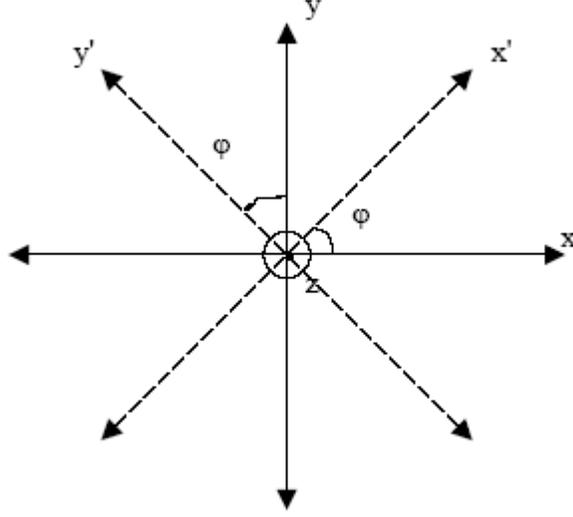
Fig. 3.3: Rotating axes x and y while keeping the z-axis fixed.

Finally, combining the effect of these two rotations, we get

$$B = \Re_\phi \Re_\theta. \tag{3.4}$$

All we know from a photon is the initial traveling direction and a series of random variables d, θ, and ϕ. When photon undergoes multiple scatterings, the effect of all these rotation matrices must be combined to find the final global position. For more illustrations, consider Fig. 3.1 and assume that a photon has undergone three scattering events so far and we have 3-tuples of random variables, $(d_0,0,0)$, (d_1,θ_1,φ_1), and (d_2,θ_2,ϕ_2). Global position of photon after these scattering events is:

$$
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ d_0 \end{bmatrix} + \begin{bmatrix} \cos\phi_1\cos\theta_1 & -\sin\phi_1 & \cos\phi_1\sin\theta_1 \\ \sin\phi_1\cos\theta_1 & \cos\phi_1 & \sin\phi_1\sin\theta_1 \\ -\sin\theta_1 & 0 & \cos\theta_1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ d_1 \end{bmatrix} +
$$
$$
\begin{bmatrix} \cos\phi_1\cos\theta_1 & -\sin\phi_1 & \cos\phi_1\sin\theta_1 \\ \sin\phi_1\cos\theta_1 & \cos\phi_1 & \sin\phi_1\sin\theta_1 \\ -\sin\theta_1 & 0 & \cos\theta_1 \end{bmatrix}\begin{bmatrix} \cos\phi_2\cos\theta_2 & -\sin\phi_2 & \cos\phi_2\sin\theta_2 \\ \sin\phi_2\cos\theta_2 & \cos\phi_2 & \sin\phi_2\sin\theta_2 \\ -\sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ d_2 \end{bmatrix}. \tag{3.5}
$$

Similarly, if a photon undergoes n scattering events, its final global position would be [26],

$$
\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = \sum_{m=0}^{n} d_m \mu_m \tag{3.6}
$$

where

23

$$\mu_m = \prod_{k=0}^{m} B_k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{3.7}$$

and

$$B_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.8}$$

$$B_k = \begin{bmatrix} \cos\phi_k \cos\theta_k & -\sin\phi_k & \cos\phi_k \sin\theta_k \\ \sin\phi_k \cos\theta_k & \cos\phi_k & \sin\phi_k \sin\theta_k \\ -\sin\theta_k & 0 & \cos\theta_k \end{bmatrix}. \tag{3.9}$$

Monte Carlo ray tracing tracks a large number of photons and records whether they have escaped from the cloud, absorbed by particles of the cloud, or arrive at the receiver plane. For the photons arriving at the receiver plane, MCRT records the position, angle of arrival, and the total traveled distance. By processing the outputs of the program, we can determine angular, spatial, and temporal dispersions of photons. For example, assuming that the total traveling distance of a photon is d_{tot} and the physical length of the channel is L_{ch}, the time delay of such a photon is:

$$t_d = \frac{d_{tot} - L_{ch}}{c}, \tag{3.10}$$

where c is the speed of light. Moreover, spatial dispersion for a photon impinging on the receiver plane at (x, y) is $r = \sqrt{x^2 + y^2}$. Also, angle of incidence of a photon hitting the receiver plane is considered as the angular dispersion.

3.4   Simulation Results

When a pulsed laser light travels through a scattering medium, photons collide with medium's particles. When the number of these collisions is large, the pulse broadens and the spatial coherence is lost. Moreover, the pulse projection on a plane perpendicular to traveling direction is no longer a small spot. Also, the intensity of the EM field is not uniform all over the beam cross-section. If we model the laser pulse with a large number of photons that are launched into the scattering medium together at the same time, we can use MCRT simulation to investigate the angular, spatial and temporal distribution of these photons on the receiver plane for various optical thickness values. To this end, photon transmission is continued until exactly 1 million photons hit the receiver plane, which is assumed to be very large. Then all the information about photon's incidence on the receiver plane is recorded. Namely, incidence point, time delay, and angle of incidence are stored in a data file. Based on these measurements, the following parameters are found.

$\overline{\cos\theta}$:   **Mean value of cosine of incidence angle on the receiver plane.**
$\overline{r}$:   **Average of photons' distance from center of the receiver plane.**
Delay Spread: **the time needed for 90% of the photons to get to the receiver plane.**
These parameters are measures of angular, spatial, and temporal dispersions, respectively. While delay spread can be normalized to the average time between two successive scatterings $(c \times k_{scat})$, $\overline{r}$ may be normalized to average distance between two successive scattering events, $\dfrac{1}{k_{scat}}$.

Fig. 3.4 and Fig. 3.5 give us a basic idea about the spatial and temporal distribution of photons at the exit plane of a cumulus cloud. While Fig. 3.4 (a) shows the spatial distribution of photons on the receiver plane for a cloud of optical thickness $\tau = 1$ (very small number of scatterings along the path), Fig. 3.5 (a) demonstrates this distribution for $\tau = 20$. When τ is small, most photons arrive at the center of the receiver plane; however, for large values of τ, a mountain of photons is formed around the center of the receiver plane. In other words, the variance of distribution increases with τ.

(a)

(b)
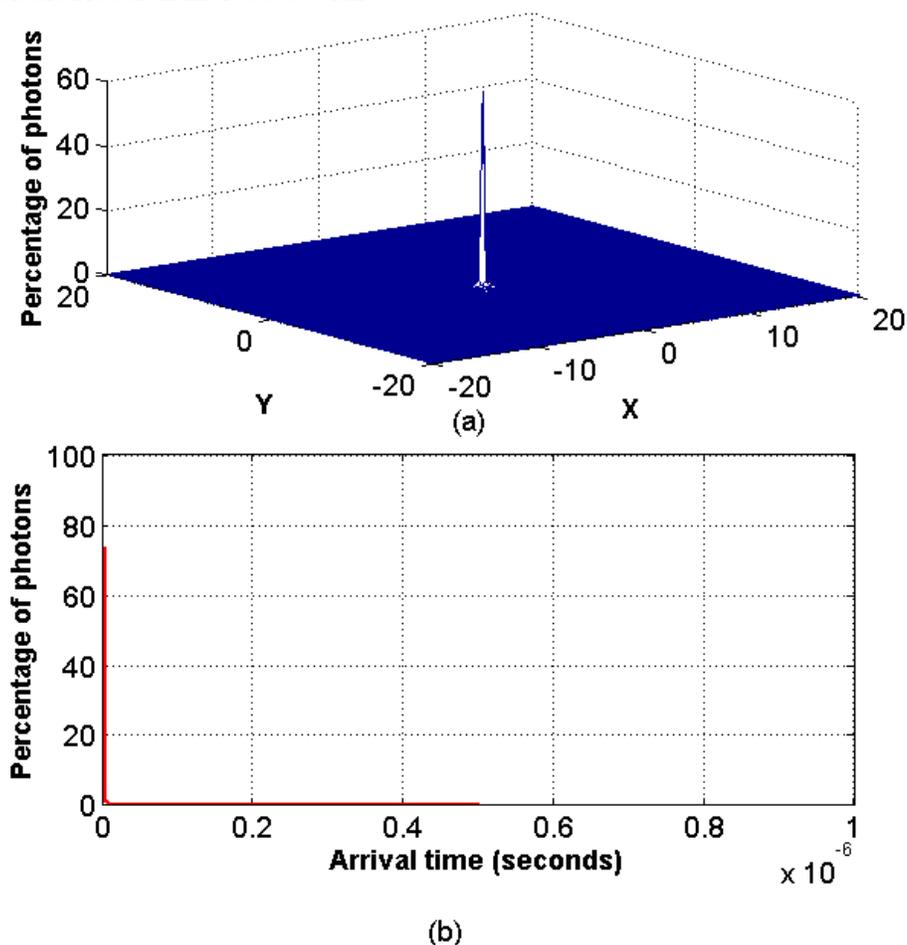
Fig. 3.4: Distribution of photons on the receiver plane for an optical thickness=1.

Temporal dispersion shows how much a pulse broadens when it travels through a scattering medium. Fig. 3.4 (b) and Fig. 3.5 (b) demonstrate the distribution of photons'

arrival times on the receiver plane. Again, we see that for τ = 1, the photons arrive approximately at the same time; however, for τ = 20, it takes much longer for 90% of photons to arrive at the receiver plane and hence we have pulse broadening. Note that, photons in MCRT are simply a symbol of energy packets. For a more quantitative analysis of angular, spatial, and temporal dispersion, we designed two scenarios so that the effect of normalization on the channel parameters can be investigated, as well. In the first scenario, it is assumed that the particulate medium is a cumulus cloud, which has the scattering coefficient of $k_{scat} = 131.8526$ km^{-1}. In this case, the optical thickness of channel is varied from 1 to 50 and the physical channel length is proportional to optical thickness ($L_{ch} = \tau D_{ave} = \dfrac{\tau}{k_{scat}}$). Then, for each of these optical thickness values (and corresponding physical channel length) the above parameters are calculated.

In the second scenario, we assume that the scattering medium has the same phase function as the cumulus cloud. However, the scattering coefficient is changed to realize different optical thickness values from 1 to 50 in a fixed physical channel length (1 km). In other words, it is assumed that transmitter and receiver are at 1 km of each other and this distance is fixed, but the density of the cloud varies and results in different optical thickness values.
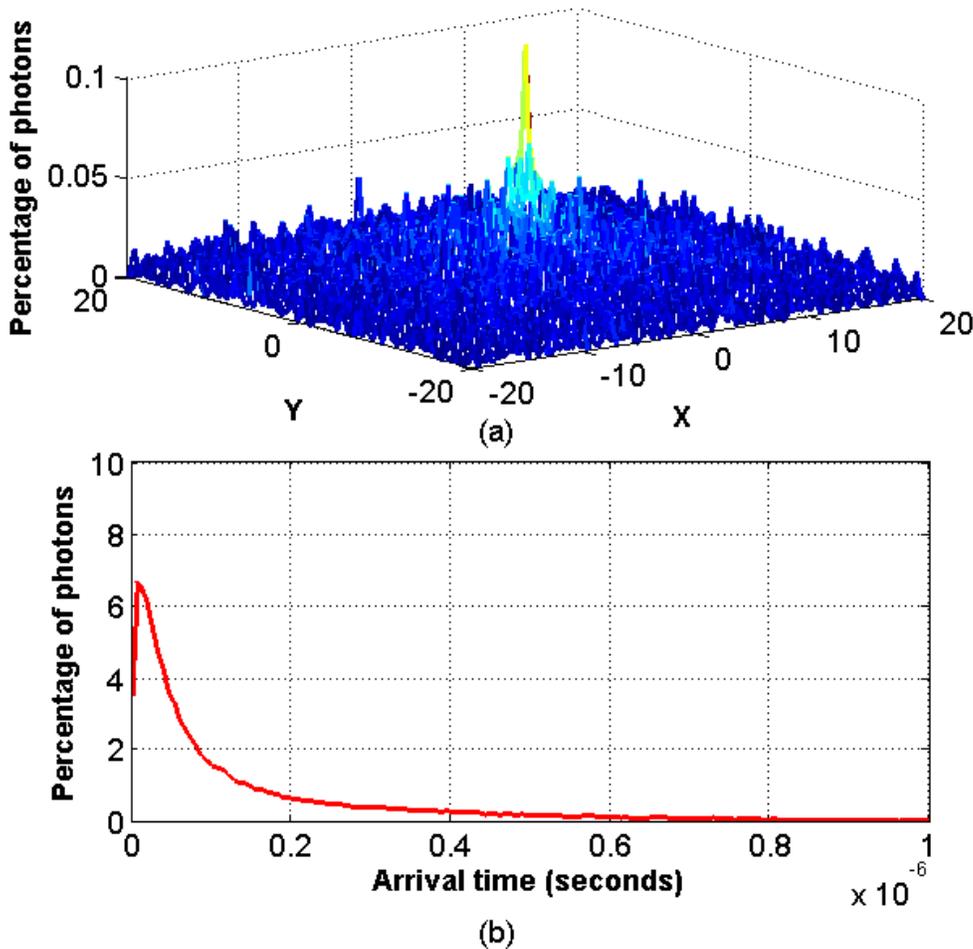
Fig. 3.5: Distribution of photons on the receiver plane for an optical thickness=20.

As mentioned before, we can normalize delay spread and $\bar{r}$ to the time and distance between two successive scattering events, respectively. Fig. 3.6 shows $\bar{r}$ (in meters) for different optical thickness values for scenarios 1 and 2. Here, one can see that in a fixed physical length channel, when the cloud becomes denser, $\bar{r}$ or spatial dispersion becomes saturated. This fact has been mentioned in [27]. Fig. 3.7 shows the normalized $\bar{r}$ for both fixed physical channel length and the case where physical channel length is proportional to optical thickness. We realize that if spatial dispersion is normalized to D_{ave}, then the results for these two scenarios are the same! From this similarity, we can understand why in Fig. 3.6 and for a fixed physical channel length, $\bar{r}$ saturates as the optical thickness value increases. This is due to the fact that for a given phase function and optical thickness, normalized $\bar{r}$ is a constant no matter what k_{scat} or D_{ave} are. However, since over a fixed physical channel, D_{ave} decreases as the optical thickness increases, $\bar{r}$ saturates at large optical thickness values. The results for the delay spread and normalized delay spread are the same as for spatial dispersion. In other words, while delay spread in units of second is different for these two scenarios and we observe the saturation of delay spread for the case of fixed physical channel length, the normalized delay spread is the same in different optical thickness values. Fig. 3.8 and Fig. 3.9 illustrate the results for delay spread in units of seconds and normalized delay spread, respectively. One can observe a similar trend for delay spread and temporal dispersion.

Now, we turn our attention to angular dispersion. We define the average cosine of incidence angle, $\overline{\cos\theta}$, as a measure of angular dispersion. We expect this value to be the same for both "fixed" and "proportional to optical thickness" physical channel lengths. Furthermore, as it is shown in Chapter 4, the beam's angular distribution converges to an isotropic one as optical thickness increases. Since, a receiver plane can catch only the forward part of this distribution and mapping from a 3D unit sphere to 2D unit circle needs a correcting factor of $\cos(\theta)$, we expect $\overline{\cos\theta}$ to converge to:

$$\overline{\cos(\theta)} = \int_0^{\pi/2} \cos(\theta)^2 \sin(\theta) d\theta = 0.6667 \tag{3.11}$$

Moreover, the variance of $\cos(\theta)$ is expected to converge to:

$$\sigma_{\cos(\theta)}^2 = \int_0^{\pi/2} \cos(\theta)^3 \sin(\theta) d\theta - \overline{\cos(\theta)}^2 = 0.0555 \tag{3.12}$$

In Chapter 4, we elaborate on angular dispersion in more details and show that (3.11) and (3.12) are indeed correct.
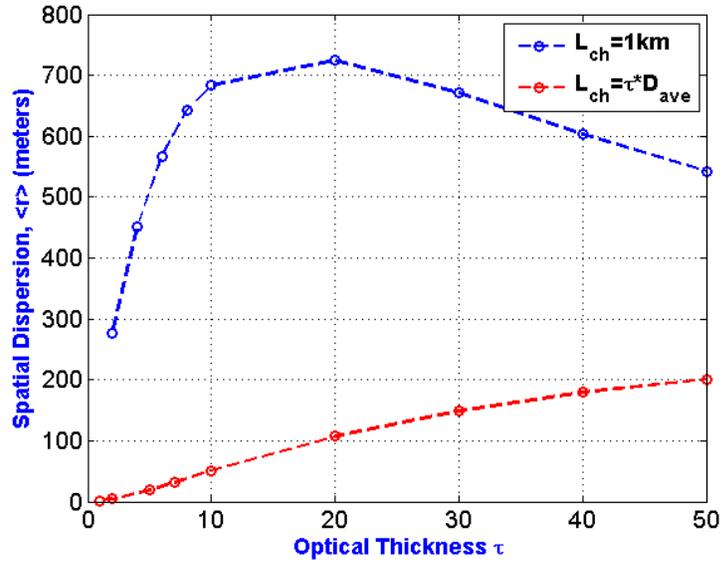
Fig. 3.6: Spatial Dispersion (meters) for "fixed" and "proportional to optical thickness" physical channel length.
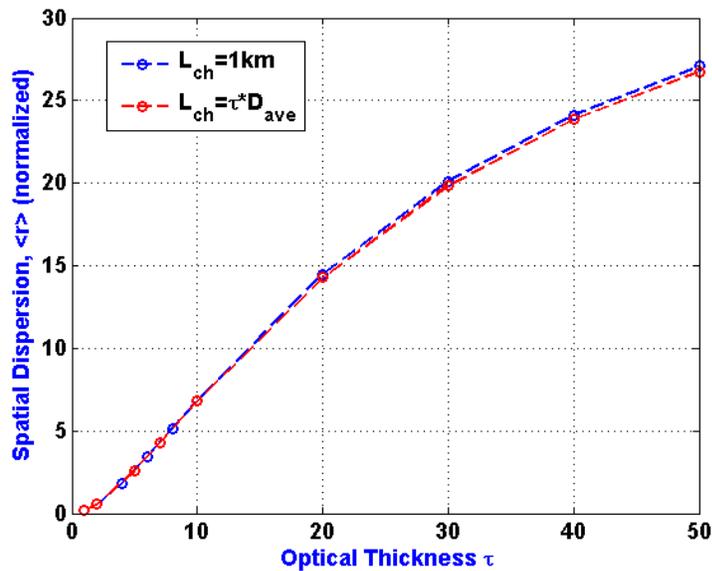
Fig. 3.7: Spatial Dispersion (normalized) for "fixed" and "proportional to optical thickness" physical channel length.

28
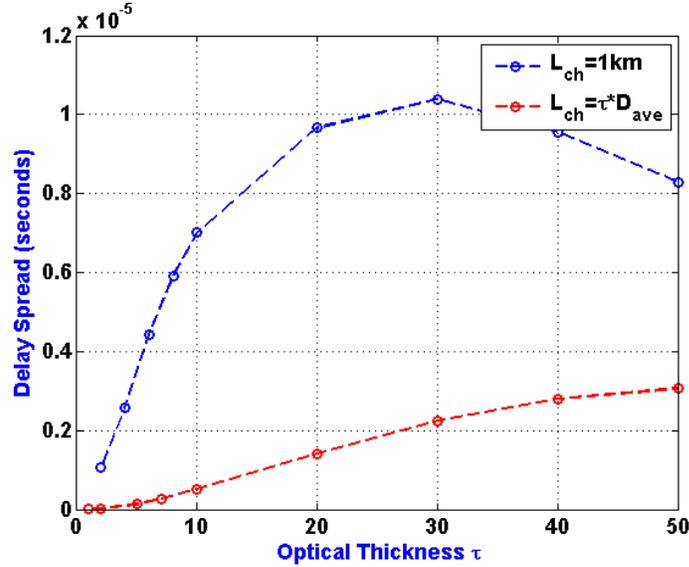
Fig. 3.8: Delay spread (seconds) for "fixed" and "proportional to optical thickness" physical channel length.
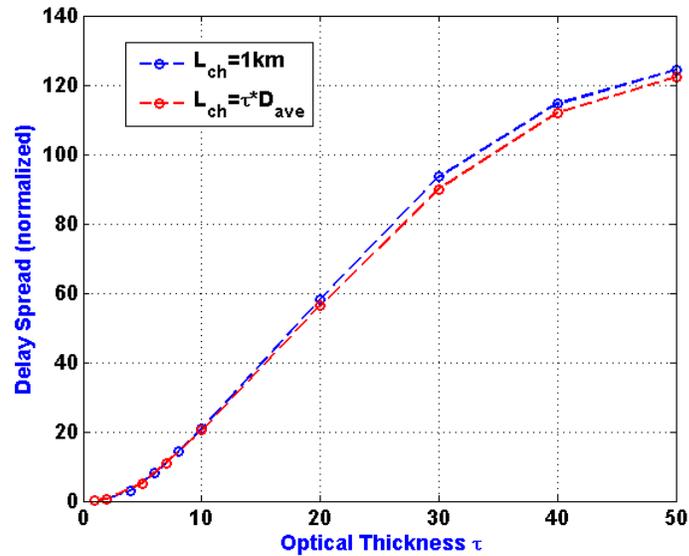
Fig. 3.9: Delay spread (normalized) for "fixed" and "proportional to optical thickness" physical channel length.

3.5   Conclusions

In this chapter, a very powerful tool for channel modeling, i.e. Monte Carlo Ray Tracing (MCRT), is introduced and described in details. It is shown that angular, spatial, and temporal dispersion increase with optical thickness. Moreover, simulation results and numbers obtained for one medium can be easily applied to any other medium through some normalization process, as dispersion is solely a function of channel's optical thickness and not its physical length. In the next chapter, we show that direct extraction of a Markov chain associated with angular distribution evolution from MCRT provides a new insight into the multi-scattering problem. Additionally, the Second Largest Eigen Modulus (SLEM) of Monte Carlo Markov

Chain is shown to be the same as the scattering phase function's asymmetric parameter. This reveals the reason behind the significance of the asymmetric parameter and the fact that Henyey-Greenstein phase function can provide the same result as that of a full Mie series phase function.

Chapter 4
Markov Chain-based Analysis

4.1   Introduction

As a collimated laser beam propagates through clouds, it spreads and after traveling some optical thickness value, it reaches a steady state of being defocused and diffused. The rate of convergence to this steady state determines feasibility of communications or imaging through various clouds of different optical thickness values.

In this chapter an analytical method is proposed to calculate angular dispersion in a scattering medium. First efforts to solve multiple scatterings problem analytically were based on the radiative transfer equation of Chandrasekhar [28], among them Dell-Imagine's [29] and Heggestad's [30] works were outstanding. However, the peaked nature of phase function in clouds and other atmospheric phenomena makes the solution of radiative transfer equation for free-space propagation a difficult one [31]. As a result, MCRT was used to solve this equation by tracking photons in a scattering medium. As mentioned before, in MCRT photons are viewed as packets of energy and not from a Quantum Mechanical perspective.

Monte-Carlo Ray Tracing (MCRT) has been used to calculate the channel parameters [27], [31], and [32]. This method requires a high computational capacity and a long execution time. In [33] authors find spatial and angular distribution by a statistical approach. However, their results are limited to the first two moments of multiple scattering. Finding angular distribution is of great importance in computer graphics, as well. In [34], an impulse response is defined for a single scatter and then the result is generalized to multiple-scattering via convolution. Since this impulse response is three dimensional, convolution process is quite cumbersome. In this chapter, angular distribution evolution of a laser beam in a multiple scattering medium is characterized by extracting the corresponding simplified Markov chain information from the MCRT algorithm. By calculating the state transition matrix, one can find the probability distribution function of the scattering angle after any number of scatterings. We show that the simplified Markov chain model produces values that are close to MCRT results and other proposed analytical methods, yet need much less time to produce the results. Since angular, spatial and temporal distributions of received power are inter-related, useful information can be extracted from angular distribution about the behavior of the entire system. While the simplified Markov chain, extracted from Monte-Carlo Markov Chain (MCMC) is limited to calculation of the angular dispersion, MCRT provides the numerical values of angular, spatial, and temporal dispersions. However, MCRT program needs much longer execution time. Also, statistical fluctuations may affect the accuracy of the results.

4.2   Markov chains

A Markov chain is a discrete-time stochastic process (Markov process) that has the Markov property. It means that the system can occupy a finite or countably infinite number of states such that the future of the process, knowing its present state does not depend on the past [35]. In other words, a stochastic process is a Markov chain if:

$$\Pr[X_n = x_n \mid X_{n-1} = x_{n-1},...,X_1 = x_1] = \Pr[X_n = x_n \mid X_{n-1} = x_{n-1}]; \ \forall \ x_1, x_2,...,x_n \in \chi. \tag{4.1}$$

A Markov chain is time-invariant if:

$$\Pr[X_n = i \mid X_{n-1} = j] = \Pr[X_2 = i \mid X_1 = j] \quad \forall n \quad \& \quad i, j \in \chi \tag{4.2}$$

That is, the conditional probability is independent of the time index. For a time invariant Markov chain,

$$p_{ij} = \Pr[X_n = i \mid X_{n-1} = j], \tag{4.3}$$

is called the state transition probability from state i to state j. The $|\chi| \times |\chi|$ matrix $P = [p_{ij}]$ is called the State Transition Matrix.

4.3 Applying Markov Chains concept to Photon Trajectories

When a photon is traveling in a scattering medium, it goes through multiple scattering until it escapes the medium's boundary. Knowing the direction and position of photon before any scattering event, we can find the probability distribution function of its direction and position before the next scattering. In other words, the most recent past determines the present, no matter what the entire history of the photon's trajectory is. Hence, it can be assumed that the trajectory of a photon is a random walk in two dimensions creating a Markov chain [31].

The MCMC, which is implemented as MCRT algorithm, specifies the kth state of a photon, i.e. photon's position and directions right before the $(k+1)^{th}$ scattering, as:

$$S_k = [\theta_k, \varphi_k, x_k, y_k, z_k], \tag{4.4}$$

where (θ_k, φ_k) represents the traveling direction in spherical coordinates and (x_k, y_k, z_k) stands for its position in Cartesian coordinates. Note that, direction can also be expressed by directional cosines (μ_x, μ_y, μ_z) in a Cartesian coordinate system [20]; however, we choose the former notation because it requires only two variables and fits better into our simplified Markov chain model. Path of each photon is determined by a series of 3-tuples (d, θ, φ), where d is an exponential random variable and stands for the distance between two successive scattering events, and θ and φ are the polar and azimuth angles of scattering with respect to previous traveling direction, respectively. MCRT calculates channel parameters such as angular, spatial, and temporal dispersions using statistics of the photons reaching the receiver plane in a post-processing stage.

In Markov chain of MCRT, variables associated with position of a photon in three-dimensional space are neither finite, nor countable. However, if we limit our attention to photons' direction in spherical coordinates, and quantize θ and φ, a finite state Markov chain is obtained. Furthermore, if it is assumed that the laser beam is traveling in a homogeneous medium, the phase function will not change from one state to the next. Thus, a Markov chain with a time-invariant state transition matrix can explain the angular distribution evolution of the laser beam, while it is traveling through a multiple scattering medium.

We claim that the angular dispersion, defined as the average cosine of the incidence angle on the receiver plane, can be calculated by modeling the photon trajectory in a 3-D space by a random walk. Moreover, this method provides us with the complete angular distribution, rather than just the moments. Also, it is superior to MCRT since due to its analytical nature, it is more tractable. Furthermore, in MCRT, a large number of photons are sent into the scattering medium in the hope of finding the distribution of photons on the receiver plane.

Hence, to account for all possible paths and angles, a large amount of processing is required that may not be feasible. While the computational complexity of MCRT is very high and the computer simulation programs need a large execution time, this analytical method can provide us with the results for angular distribution through simple matrix multiplications.

Suppose a photon is traveling in the θ_{k-1} direction with respect to z axis before k^{th} scattering. If this photon collides with a cloud particle, its propagation direction changes in space. Fig. 4.1 illustrates the geometry of the problem. The change of direction is described by the phase function. However, phase function provides PDF of cosine of deflection angle with respect to initial traveling direction, while we are interested in a traveling direction with respect to the global z axis. In other words, the phase function provides us with the distribution of $\cos(\theta)$, while we want the distribution of θ_k, as in Fig. 4.1. The term $\cos(\theta_k)$ is related to $\cos(\theta_{k-1})$ via:

$$\cos\theta_k = \cos\theta_{k-1}\cos\theta - \cos\varphi\sin\theta_{k-1}\sin\theta, \tag{4.5}$$
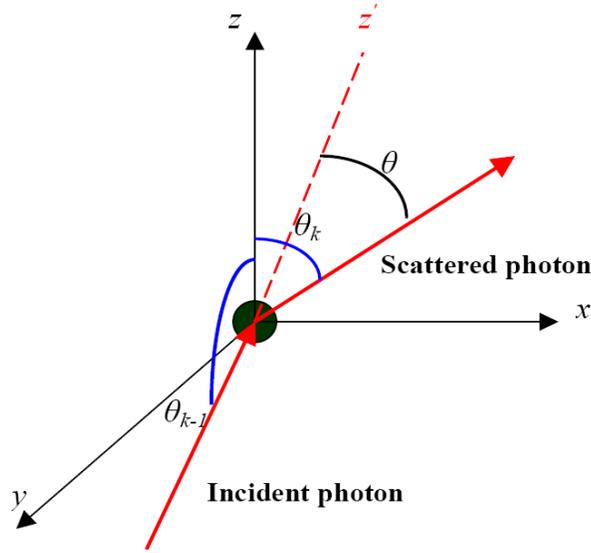
Fig. 4.1: Relationship between θ_{k-1}, θ and θ_k.

where θ and φ are polar and azimuth scattering angles, respectively. While φ is uniformly distributed in $[0, 2\pi]$, the distribution of θ is given by $P(\theta) \cdot \dfrac{\sin(\theta)}{2}$. To calculate the PDF of θ_k, first we determine its cumulative distribution function (CDF).

$$\Pr(\theta_k < \xi) = \Pr(\cos\theta_k > \cos\xi) =$$
$$\Pr(\cos\theta_{k-1}\cos\theta - \cos\varphi\sin\theta_{k-1}\sin\theta > \cos\xi) \tag{4.6}$$

Hence, we can write

$$\Pr\left(\cos\varphi < \frac{\cos\xi - \cos\theta_{k-1}\cos\theta}{-\sin\theta\sin\theta_{k-1}}\right) =$$
$$\Pr\left(\varphi > \cos^{-1}\left(\frac{\cos\xi - \cos\theta_{k-1}\cos\theta}{-\sin\theta\sin\theta_{k-1}}\right)\right) \tag{4.7}$$

Finally, the cumulative distribution can be expressed as:

$$\Pr(\theta_k < \xi) = \int_0^\pi f(\theta) \int_{\cos^{-1}\left(\frac{\cos\xi - \cos\theta_{k-1}\cos\theta}{-\sin\theta\sin\theta_{k-1}}\right)}^{2\pi} f(\varphi)d\varphi d\theta$$

$$= \int_0^\pi P(\theta)\sin\theta\left(1 - (2\pi)^{-1}\cos^{-1}\left(\frac{\cos\xi - \cos\theta_{k-1}\cos\theta}{-\sin\theta\sin\theta_{k-1}}\right)\right)d\theta$$

(4.8)

If we differentiate the last equation, we obtain the PDF of θ_k. It is also possible to directly generate PDF of θ_k, by rewriting as:

$$\cos(\theta_k) = \cos(\varphi/2)^2\cos(\theta_{k-1} + \theta)$$
$$+ \sin(\varphi/2)^2\cos(\theta_{k-1} - \theta)$$

(4.9)

Then, the PDF of θ_k can be found by adding up the probabilities associated with all values of θ and φ that give rise to a certain θ_k, given θ_{k-1}. Now, we can form a matrix with its rows corresponding to incident angle and its column corresponding to scattering angle. This matrix can be used as the state transition matrix of a Markov process. (4.10) illustrates this matrix.

$$P = \begin{bmatrix} P(\theta_k = 0 \mid \theta_{k-1} = 0) & \dots & P(\theta_k = \pi \mid \theta_{k-1} = 0) \\ . & & . \\ . & \dots & . \\ . & & . \\ P(\theta_k = 0 \mid \theta_{k-1} = \pi) & \dots & P(\theta_k = \pi \mid \theta_{k-1} = \pi) \end{bmatrix}$$

(4.10)

If one wishes to have the distribution of photon's direction after k^{th} scattering, knowing that it was initially traveling in the z direction, one has to calculate

$$P_k(\theta) = (P^K)^T \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

(4.11)

First row of the state transition matrix is *P(θ)sin(θ)/2*. Hence, we can find the phase function of *k* scattering events from the first row of the kth power of the state transition matrix, simply by taking out the *sin(θ)/2* factor.

Fig.4.2 shows the state transition matrix for the cumulus cloud at a wavelength of *1.55 μm*. We have calculated the state transition matrix with the resolution of $\pi/300$. In other words, *P* is a 300×300 matrix. We observed that by increasing the resolution to $\pi/1000$, our simulations result does not change.
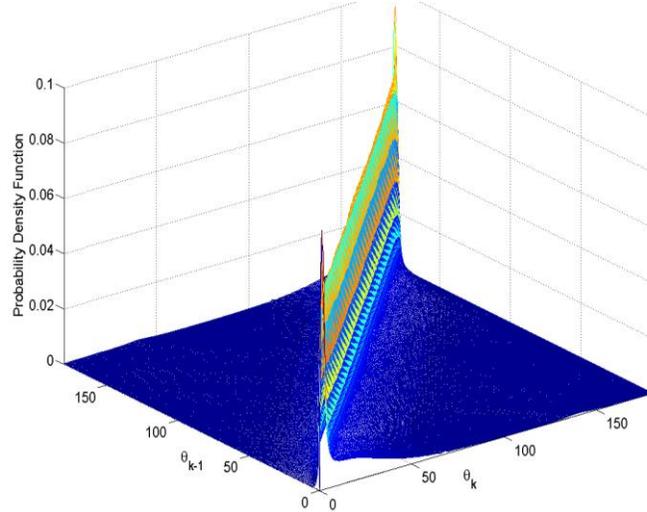
Fig.4.2: State transition matrix of cumulus clouds at a wavelength of $\lambda = 1.55\mu m$.

The forward scattering property of cumulus cloud is clear from Fig.4.2. That is, the state transition matrix is very close to an identity matrix. Fig. 4.3 shows P^{15} which clarifies transition probabilities at an optical thickness 15.
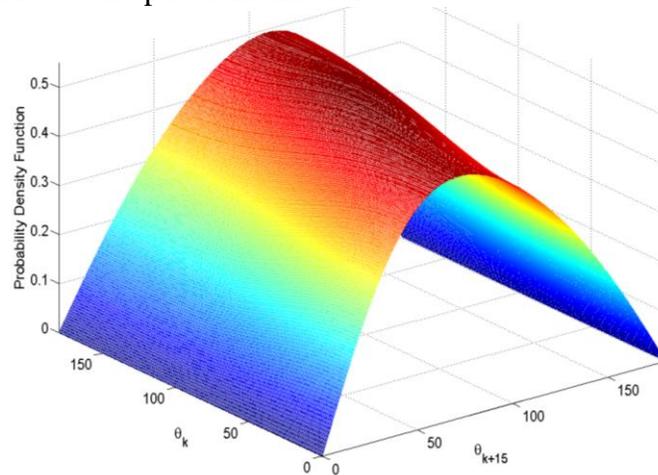
Fig. 4.3: State transition matrix of cumulus cloud, raised to 15th power, at a wavelength of $\lambda = 1.55\mu m$.

Here, the optical thickness 15 is chosen since for this value, there are less line-of-sight (LOS) photons, and most of the photons that may reach the receiver have gone through multiple-scatterings, hence, this effect can be illustrated. From Fig. 4.3 one can see that after 15 scattering events, the phase surface is very close to one shaped as *sin(θ)/2*. This corresponds to an isotropic scattering. Fig. 4.4 shows a side view of the phase surface. One can observe that different rows of the matrix are very close but, not as yet identical. Fig. 4.5 shows P^{50}, that is, the phase surface after 50 scattering events. This figure clearly illustrates isotropic radiation after 50 scattering events. Fig. 4.6 shows a side view of the phase surface. It can be inferred that after 50 scattering events, no matter what the initial incident angle value, the scattering angle is distributed as *sin(θ)/2*. This is consistent with Bucher's [27] observation of uniform brightness of the cloud bottom.
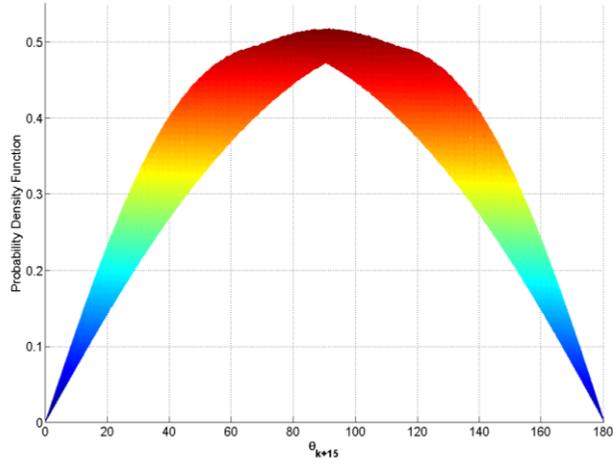
Fig. 4.4: Side view of the state transition matrix, raised to 15th power.
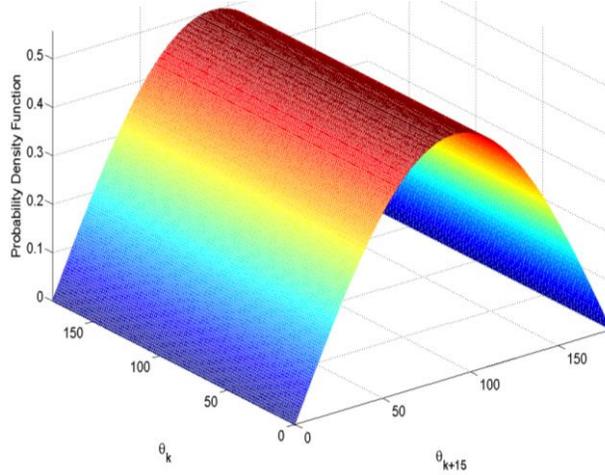
Fig. 4.5: State transition matrix of cumulus cloud, raised to 50th power, at a wavelength $= 1.55 \mu m$.
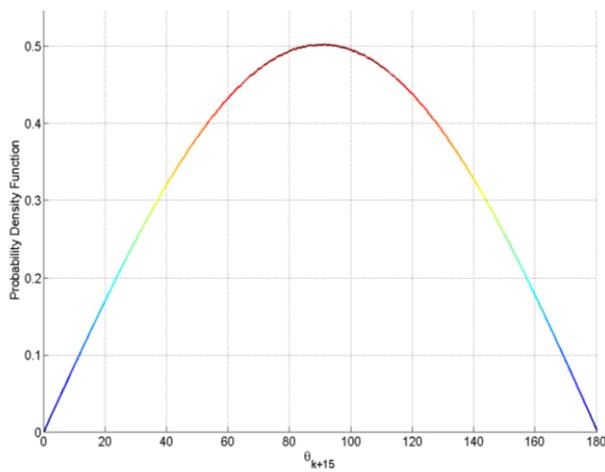
Fig. 4.6: Side view of the state transition matrix, raised to 50th power.

4.4　　　*Comparison of Results*

As mentioned earlier, application of Markov chain in MCRT cloud modeling is a shortcut for calculating angular distribution of energy in space in any arbitrary optical thickness. In this section, we compare the results of our Markov chain model with those by MCRT and another analytical method, which we call the moment technique [33].

To make a fair comparison, we note that while both MCRT and Markov chain model produce a complete distribution, the moment technique only comes up with the first two moments. Furthermore, MCRT can provide us with the distribution of photons on the receiver plane. However, both the Markov chain model and the moment technique generate angular distribution in a three-dimensional space, i.e. on a sphere. Hence, to compare the Markov chain model with MCRT, we should only consider the forward part of the distribution, i.e., $0 < \theta < \pi/2$. Moreover, we should take into account the projected area correction factor of $\cos(\theta)$ for mapping from 3-D distribution onto a sphere to 2-D distribution on the receiver plane [27]. Given a photon at the cloud exit plane has scattered k times over a cloud length L, its angular distribution would be $P_k(\theta)$ as in (4.11). To calculate the unconditional angular distribution, we note that the probability that a photon undergoes exactly k scatterings over L is Poisson distributed with a mean τ, where τ is the optical thickness, as defined earlier. That is:

$$P(K(\tau) = k) = \frac{\tau^k}{k!} e^{-\tau}. \tag{4.12}$$

Hence, the angular distribution for this optical thickness value is:

$$P(\theta) = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} e^{-\tau} . P_k(\theta) \tag{4.13}$$

Using (4.13) , angular distribution is calculated for optical thickness values of 1 to 15. Fig. 4.7 shows the cumulative distribution function (CDF) of incident angle obtained from Markov chain model and MCRT, where we have applied the above-mentioned measures in order to make the comparison fair. From Fig. 4.7, one can see that CDF curves obtained from Markov chain model and MCRT are quite close in numerical values. Furthermore, as optical thickness value increases, angular dispersion increases.
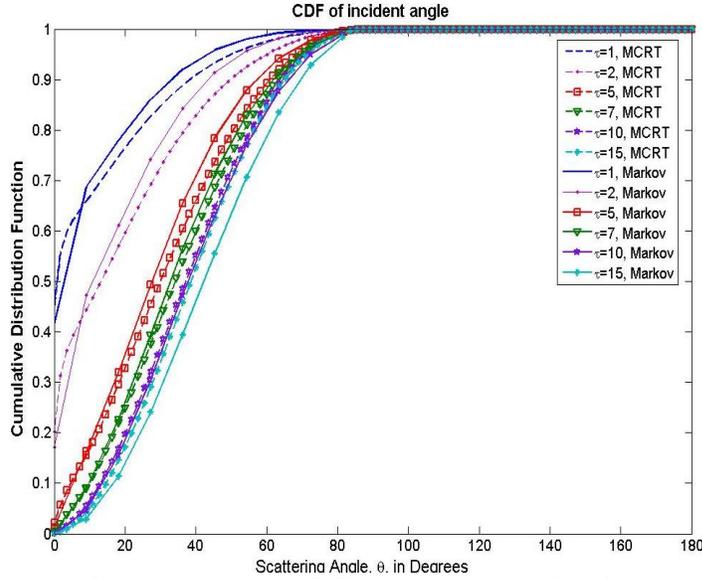
Fig. 4.7: Comparison of *CDF* curves for angular distribution.

One important measure of angular dispersion is the average cosine of incident angle, $\overline{\cos(\theta)}$. Variance of incident angle cosine is also considered as a measure of angular dispersion in the literature. The moment technique provides us with the first two moments of *cos(θ)* in a three-dimensional space. However, MCRT provides the complete distribution on the receiver plane. Then, one can find the first two moments of $\cos(\theta)$ using this distribution. The Markov chain model, in its original shape, produces the same mean and variance as those of the moment technique. However, as mentioned earlier, it can be truncated and modified to provide the distribution on the receiver plane. Fig. 4.8 and Fig. 4.9 show the mean and variance of *cos(θ)* for MCRT, truncated Markov, original Markov, and moment technique (labeled 'Moments') for different optical thickness values.
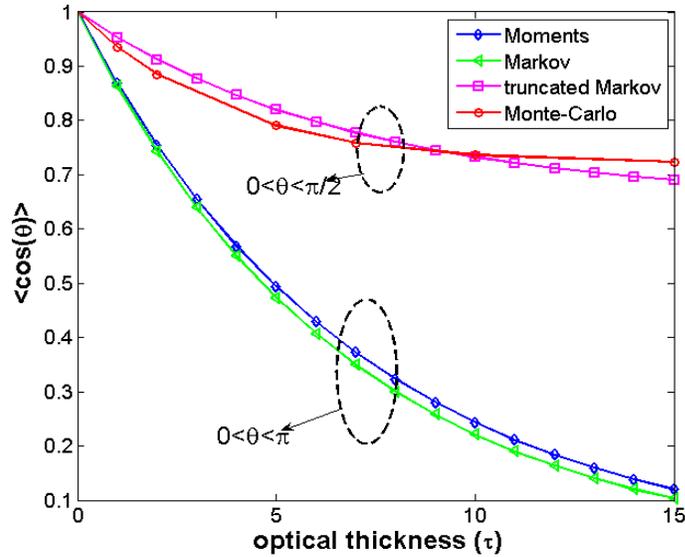
Fig. 4.8: Comparison of $\overline{\cos(\theta)}$ for different methods.
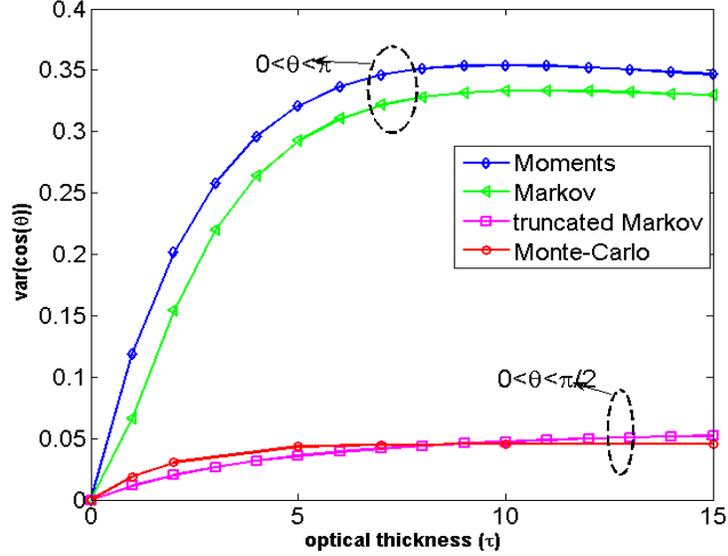
Fig. 4.9: Comparison of $\mathrm{var}(\cos(\theta))$ for different methods.

It is clear from these figures that the variance of the cosine of incident angle increases with optical thickness, whereas the mean decreases. Moreover, saturated values of mean and variance of incident angle cosine are as predicted in chapter 3. This implies that for small optical thickness values, a receiver of small FOV suffices to collect the required power; however, for large optical thickness values, the energy is almost uniformly distributed in a 3-D space. Saturation of variance curves in Fig. 4.8 and Fig. 4.9 suggests that density evolution of angular distribution merges to the steady-state and the Markov chain converges to the equilibrium distribution. It is also evident from these figures that our Markov chain model is consistent with both MCRT and moment technique. Finally, the time required for each method to generate the results in different optical thickness values is tabulated in Table 4.1. Computer programs were executed on a 3.40 GHz Pentium-4 CPU, with 2.00 GB of RAM. It is clear from this table that the computation time for MCRT increases with the optical thickness value. However, neither the Markov chain model, nor the moment techniques show much variability in the computation times. We also note that the execution time required for the Markov chain model is significantly less than MCRT but much more than moment technique. However, moment technique gives only the first two moments, and not the complete distribution. Thus, the Markov chain model is able to demonstrate a more complete picture.

Table 4.1: Time required for each method to generate the results in different optical thickness values.

Optical Thickness	MCRT	Markov	Moments
1	656 s	33 s	0.06 s
5	767 s	33 s	0.06 s
10	948 s	33 s	0.06 s
15	1212 s	33 s	0.06 s

4.5 Eigen Analysis

State transition matrix, P has no zero entries, and hence is regular. In other words, it is possible to go to all the states from any arbitrary state. It is a well known fact that for a regular Markov chain, as n approaches to infinity, $P^n \rightarrow \Pi$, where Π is a matrix of the form *[v, v, ..., v]*, with *v* being a constant vector. From the previous section, we see that this is true about the state transition matrix of our Markov chain P, and as n increases, all the rows of the state transition become identical and proportional to *sin(θ)/2*. Now, the question is whether it is possible to predict v, and thus Π, without using the limits and the answer is affirmative. In fact, Π satisfies the equation:

$$\Pi P = \Pi \tag{4.14}$$

Notice that, from Perron-Frobenius theorem , v is the first left eigenvector of P, corresponding to the unique largest eigenvalue, $\lambda_0=1$. By examining the first left eigenvector of matrix P, we realize that it is proportional to $\sin(\theta)/2$. Hence, we could have predicted the diffuse behavior of light at the bottom of the cloud, just by looking at the left eigen vector of the state transition matrix, from the beginning.

Convergence of P^k elements means that it becomes more and more difficult to guess k, from p_{ij}^k (the element in the ith row and the jth column of P^k). That is, the chain forgets the length of its history [36]. The fact that limit of P^k has identical rows suggests that the Markov chain forgets the initial position.

When a Markov chain converges to this steady-state, traveling direction of photons becomes rather isotropic, as opposed to forward-scatter condition in initial steps. This suggests that the laser beam is spatially diffused and no matter what the initial traveling direction is, photons escape almost uniformly from all boundaries of cloud. In this case, spatial confinement of transmitted energy is no longer preserved and space loss (mostly attributed to scattering) is rather large.

In some applications, due to eye safety regulations, there is restriction on increasing the transmit power level beyond a certain maximum. Under such circumstances, receiver may not see much forward scatter (snake) photons and must resort to only the LOS (ballistic) photons, that is, the non-diffused part of the intensity, which is attenuated according to Beer-Lambert law as:

$$I_{coh} = I_0 e^{-\tau} \tag{4.15}$$

Needless to say, by using ultra-short laser pulses and cascade amplifiers, this component can be amplified, significantly. In (4.15), τ is the optical thickness of cloud, as defined earlier.

The number of steps required for the Markov chain to converge to the equilibrium state is of great importance since it determines the depth up to which the laser beam can penetrate before becoming spatially diffused. From (4.5), one can see that:

$$\overline{\cos(\theta_k)} = \overline{\cos(\theta_{k-1})} \times \overline{\cos(\theta)} = g\overline{\cos(\theta_{k-1})} = g^k \qquad (4.16)$$

In other words, the asymmetric parameter, g determines the convergence rate of Markov chain. On the other hand, from Markov chain theory [37], we know that the Second Largest Eigenvalue Modulus (SLEM) determines the "Mixing Rate" of a Markov chain. In other words, the smaller the SLEM is, the faster the spatial Memory-loss happens. Hence, "Mixing-Time" of a Markov chain is given by:

$$T = \frac{1}{\log\left(\dfrac{1}{\lambda_*}\right)} \qquad (4.17)$$

where T is the number of steps over which deviation from equilibrium state decreases by a factor e, and λ_* is the SLEM. By examining the eigenvalues of P (state transition matrix of cumulus cloud) we realize that the second largest eigenvalue is indeed the asymmetric parameter g.

Table 4.2 lists the SLEM for Markov chains associated with HG phase functions of different asymmetric parameters as well as full Mie series phase function for different types of clouds. This table also contains the mixing time and the average distance between two successive scattering events for each cloud. Average distance, D_{ave}, between two successive scattering events is used to convert optical thickness of a specific scattering medium to its physical thickness and vice versa. It also corresponds to the inverse of scattering coefficient, expressed in km^{-1}, which can be obtained from substituting clouds particle size distribution in equations extracted from Mie theory for poly-dispersed extinction coefficient [15], [38]. From Table 4.2 we see that the second largest eigenvalue is equal to g for all of these phase functions. That is why asymmetric parameter is so important in calculating the moments of multiple scattering and channel parameters [33], [27]. Since angular, spatial and temporal distributions of energy are inter-related, isotropic angular distribution suggests that spatial distribution in transverse coordinates (x and y directions) is Gaussian. Furthermore, spatial memory-loss of Markov chain associated with angular distribution after several scattering events implies that the entire Monte-Carlo Markov chain (MCMC) has converged to the equilibrium state. This suggests that the pulse is so broadened in space and time that use of equalization on the pulse might be difficult, unless there is no constraint on the optical transmitted peak power level.

Examining parameters of Table 4.2, we realize that mixing time is rather short for clouds and haze and a large link margin is necessary, in order to have reliable imaging through clouds of optical thickness values much larger than 10 to 15. However, this rather small optical thickness may translate into a long physical thickness for clouds such as thin Cirrus due to the long average distance between two successive scatterings for these clouds [15].

Note that, all the results presented in this paper assume a homogeneous body of clouds. In reality, as a waveform moves through clouds, the dynamics are far more rapidly varying. Hence, the predictions in this chapter are on the conservative side.

Table 4.2: Asymmetric parameter and SLEM for different clouds at a wavelength of 1.55 μm

Cloud	SLEM	g	Mixing Time	D_{ave}
Thin Cirrus	0.82	0.82	5	11.3 km
Cirrus	0.87	0.87	7	984 m
Alto Stratus	0.83	0.83	5.4	10.5 m
Nimbo Stratus	0.85	0.85	6.2	12.3 m
Stratu Columbus	0.82	0.82	5	26.5 m
Stratus	0.83	0.83	5.4	17.5 m
Cumulus	0.85	0.85	6.2	7.5 m
HG(g=0.85)	0.85	0.85	6.2	N/A
HG (g=0.95)	0.95	0.95	19.5	N/A
Low Altitude Haze	0.73	0.73	3.2	817.8 m
Medium Altitude Haze	0.78	0.78	4	238.8 m

4.6   Conclusions

To answer the fundamental question of feasibility of active optical imaging in a scattering medium, one has to accurately model atmospheric optical channel and estimate parameters such as angular, spatial, and temporal dispersion. As a collimated laser beam propagates through aerosols and cloud particles, it spreads and after traveling some optical thickness reaches the steady-state of being nearly diffused. The rate of convergence to this steady-state determines the feasibility of imaging through various types of clouds and aerosols with different optical thickness values. Average distance between two successive scattering events is used to convert optical thickness of a specific scattering medium into its physical thickness and vice versa.

By directly applying Markov chain model to angular distribution evolution of laser beam in a scattering medium, and considering the fact that angular, spatial and temporal distributions of energy are inter-related, the mixing rate of Monte-Carlo Markov Chain (MCMC) is found for different types of scattering media. Mixing time is referred to the number of steps over which deviation from equilibrium state decreases by a factor e. This indicates the rate by which the energy distribution approaches being spatially isotropic.

However, it may take many more scatterings for the temporal coherence of the laser beam to be lost.

Examining parameters of Table 4.2, we realize that mixing time is rather short for clouds and haze. That is, when coherent light is traveling through clouds, its spatial coherence rapidly degrades. However, this may still translate into a long physical thickness for clouds such as thin Cirrus due to the rather long average distance between two successive scatterings in these clouds [15]. Also, all these results assume a homogeneous body of clouds that do not exist in the atmosphere. Therefore, experimentally, one may succeed transmitting over longer lengths of various clouds.

Chapter 5
Atmospheric Turbulence and Scintillation

5.1 Introduction

Optical wave propagation through atmosphere is affected by temperature fluctuations related to the sun's heating of the earth and the turbulent motion of the air pockets due to winds and convection [39]. Hence, all the optical systems, that encounter atmosphere as a propagation medium, experience major performance degradations. In Free Space Optical (FSO) communications systems, turbulence manifests itself through beam-wander, scintillation, and flat fading. In optical imaging systems on the other hand, depending on the exposure time, turbulence gives rise to Point Spread Function (PSF) broadening and speckled appearance, limiting the ability to measure high resolution information [6].

In imaging through atmosphere one should consider the dispersive effect of scattering, the turbulence-induced beam spreading, intensity fluctuations, and wave-front distortion. All these phenomena broaden and distort the Point Spread Function (PSF) which serves as the impulse response of the imaging system. In this monthly report, distortion of the PSF is investigated and different methods of simulating the wave-front distortion are presented. More specifically, a phase screen is used to simulate the phase perturbations introduced into the wave-front as a result of scintillation.

Turbulence induced fluctuations can distort both amplitude and phase of the wave-front incident on the receiving aperture. Fig. 5.1 shows the diffraction pattern of a plane wave generated in the pupil plane of a circular aperture, assuming no turbulence. We can see that the diffraction pattern is the same as Airy pattern. In Fig. 5.2, we see the diffraction pattern of a plane wave in presence of weak turbulence in pupil plane of the same lens. One can observe the broadening of the diffraction pattern in presence of turbulence. In the case of moderate turbulence the diffraction pattern may also shift from the center of the pupil plane. Fig. 5.3 shows the diffraction pattern for moderate turbulence.

Attributes of turbulent atmosphere are usually simulated using thin phase screens which perturb the phase of a propagating wave-front in accordance with either Kolmogorov or von Kàrmàn model. FFT-based methods are among the most popular methods of generating phase screens. However, these techniques are limited in that they do not accurately reproduce the low spatial frequency characteristics of Kolmogorov turbulence such as wave-front tilt. There are several techniques to extend the FFT based phase screens to include the low spectral frequency components, namely; sub-harmonics and random mid-point displacement methods.
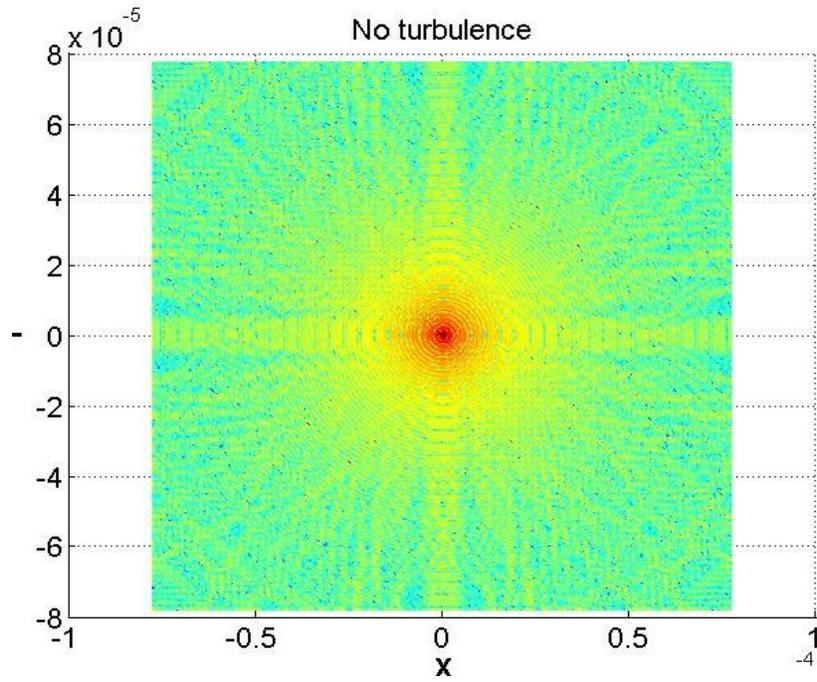
Fig. 5.1 Airy pattern formed in pupil plane of a circular lens in absence of turbulence.
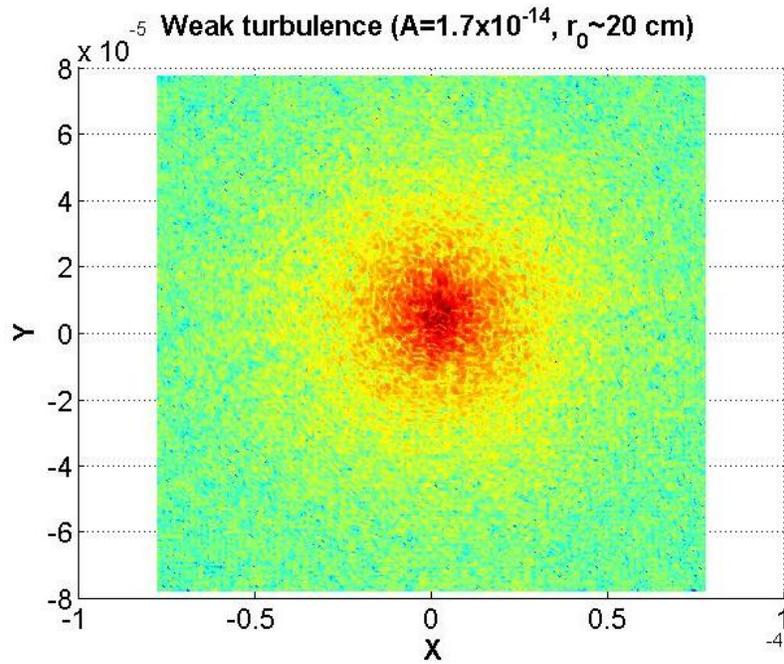
Fig. 5.2 Diffraction pattern in pupil plane of a circular lens in presence of weak turbulence.
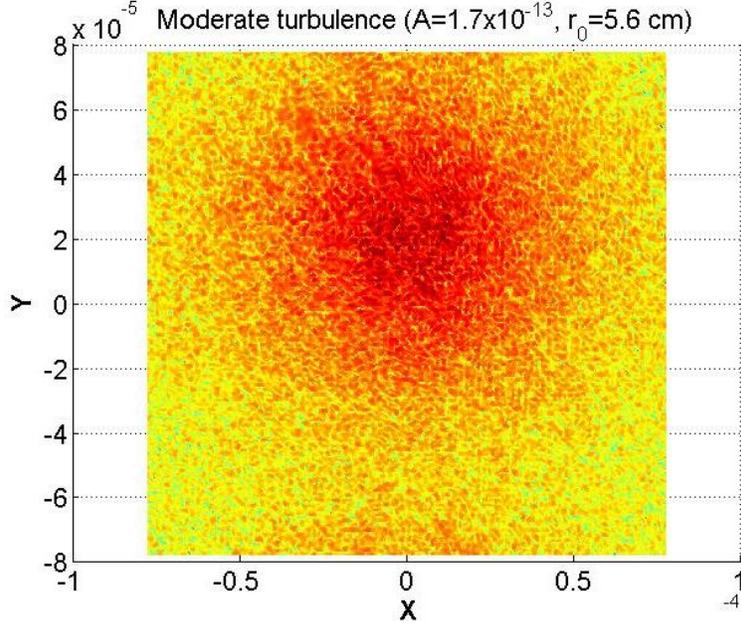
Fig. 5.3 Diffraction pattern in pupil plane of a circular lens in presence of moderate turbulence.

In order to analyze, formulate, and simulate the complex nature of atmospheric turbulence, we need simplifying models and approximations. In the subsequent sections, we talk about Kolmogorov theory of turbulence and Rytov approximation, which model the wave-front distortions. Then we elaborate on the thin phase screen as a tool for simulating the wave-front distortion. Next, different methods of generating phase screens, such as random mid-point displacement algorithm, sub-harmonics method, and Zernike polynomials technique are described. In the next chapters, these tools are used to model the propagation medium and the resulting degradations in system performance.

5.2   Kolmogorov Theory of Turbulence and Rytov Approximation

Temperature variations and wind give rise to spatial and temporal variations in the refractive index of the atmosphere. As a result of these variations, a propagating wave experiences random spatial and temporal fluctuations in the optical path length (OPL), through a continuum of altitudes, h. These OPL variations are denoted by $n_1(\vec{x}, h)\Delta h$, where $n_1(\vec{x}, h)$ represents the fluctuation of the index of refraction about the mean value, Δh is the thickness of a thin layer of the atmosphere, and $\vec{x}$ is a two dimensional vector location [6]. Hence, the phase change due to such propagation through a thin layer of atmosphere is

$$\psi(\vec{x}, h) = kn_1(\vec{x}, h)\Delta h, \qquad (5.1)$$

where $k = \dfrac{2\pi}{\lambda}$ is the wave number of light with wavelength, λ. Though all atmospheric turbulence-induced wave-front aberrations start as phase distortion, propagation of the distorted wave-front along with refraction can cause different portions of wave to interfere with one another and give rise to amplitude variations over long distances.

46

Optical turbulence can be characterized by three parameters: inner scale l_0 of turbulent eddies, outer scale L_0 of turbulent eddies, and the structure parameter of the refractive index fluctuations C_n^2 [39]. The well-known Hufnagle-Valley equation models the profile of C_n^2 as:

$$C_n^2(h) = 0.00594(v/27)^2(10^{-5}h)^{10}\exp(-h/1000) + $$
$$2.7 \times 10^{-6}\exp(-h/1500) + A\exp(-h/100), \tag{5.2}$$

where h is the altitude in meters (m), v is the rms wind speed in meters per second (m/s), and A is the nominal value of $C_n^2(0)$. Fig. 5.4 demonstrates the profile of refractive index fluctuations structure parameter. C_n^2 is a measure of the strength of turbulence. Variations in the refractive index, give rise to deformation of the wave-front (iso-phase plane), which cause phase perturbations in the receiver plane. Hence, under near field conditions, we only expect to observe phase perturbations. Under far field conditions, bending of the optical rays caused by refraction, along with the subsequent propagations to the pupil plane make different portions of the wave interfere with one another. The amplitude variations are caused by these interference elements. In other words, phase perturbations evolve into both amplitude and phase fluctuations as a result of propagation [6].
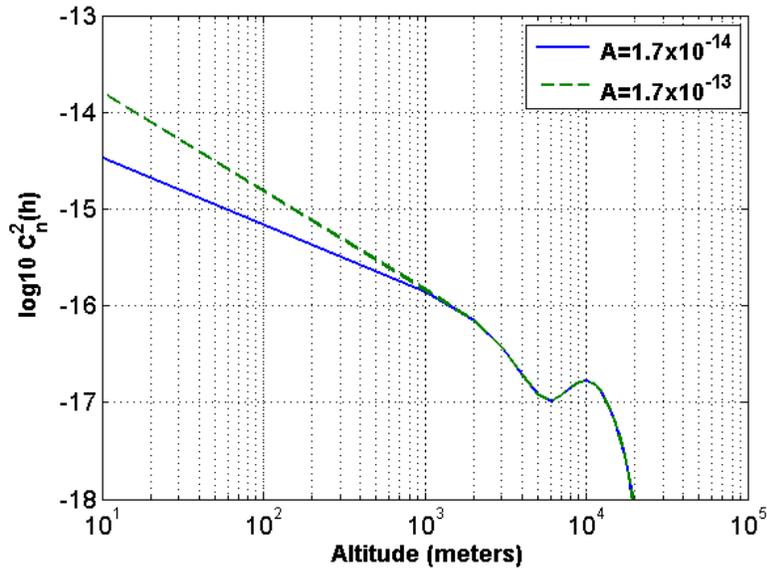
Fig. 5.4: $C_n^2(h)$ profile associated with H-V model as a function of altitude.

According to Rytov approximation, atmospheric turbulence can be modeled as a complex multiplicative factor at each point on the wave-front [39]. For example, the optical field at a distance L from the transmitter is given by:

$$U(r,L) = U_0(r,L)\exp[\psi(r,L)] = U_0(r,L)\exp[X(r,L)+iS(r,L)] \tag{5.3}$$

where $X(r,L)$ represents amplitude fluctuations and $S(r,L)$ represents phase variations. Using central limit theorem, for long propagation distance through turbulence, it can be

shown that X and S are homogeneous, isotropic and independent Gaussian random variables [5]. According to (5.3) both amplitude and intensity are log-normally distributed at each point on pupil plane. Under weak turbulence conditions, the log-irradiance variance $\sigma_{\ln I}^2 = 4\sigma_x^2$ (σ_x^2 is the log-amplitude variance) is approximately equal to the normalized variance of irradiance or the scintillation index, i.e., $\sigma_I^2 \approx \sigma_{\ln I}^2$. The scintillation indices of plane and spherical waves are given by:

$$\sigma_I^2 \mid_{plane} = 2.25 k^{7/6} \int_0^L C_n^2(h)(1 - h/L)^{5/6} \, dh \tag{5.4}$$

$$\sigma_I^2 \mid_{spherical} = 2.25 k^{7/6} \int_0^L C_n^2(h) h^{5/6} (1 - h/L)^{5/6} \, dh \tag{5.5}$$

Furthermore, to make sure that the average optical field amplitude is neither attenuated nor amplified, the mean value of log-irradiance is set to $-\sigma_I^2/2$. Hence, the PDF of irradiance can be expressed as:

$$f_I(I) = \frac{1}{I\sqrt{2\pi\sigma_I^2}} \exp\left\{ -\frac{(\ln(I) + \sigma_I^2/2)^2}{2\sigma_I^2} \right\} \tag{5.6}$$

Phase perturbation is usually approximated by a thin phase screen, the power spectrum of which is given by the Kolmogorov or von Kàrmàn model. According to the Kolmogorov theory, assuming isotropic and homogenous turbulence, the power spectral density of refractive index fluctuations can be expressed as:

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3}, \ 1/L_0 < \kappa < 1/l_0, \tag{5.7}$$

where κ is the spatial wave number. Fig. 5.5 illustrates the normalized Kolmogorov spectrum along with von Kármán spectrum, which is developed to consider the inner and outer scale effects, as well as a more precise model based on analytic approximation to the Hill spectrum [39].
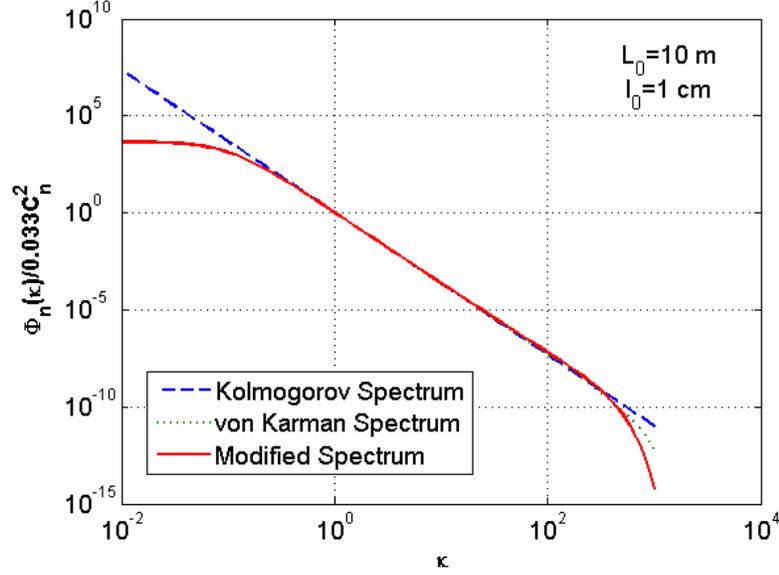
Fig. 5.5: Spectral models of refractive-index fluctuations.

Given the Kolmogorov spectrum for refractive index variations, the power spectrum of phase fluctuations is represented as:

$$\Phi_p(k) = 0.023 r_0^{-5/3} \, |\kappa|^{-11/3}, 1/L_0 < \kappa < 1/l_0 \tag{5.8}$$

where r_0 is the atmospheric coherence length (Fried parameter) and is approximated as:

$$r_0 = 0.185 \left[\frac{4\pi^2}{k^2 \int_0^L C_n^2(h)(1-h/L)^{5/3} \, dh} \right]^{3/5} \tag{5.9}$$

This parameter is a measure of coherence radius of the optical field. As turbulence becomes stronger, r_0 decreases. Since coherence length of the field is proportional to r_0, if two receivers are r_0 apart, they observe virtually independent versions of the signal. The phase structure function can be defined as:

$$D_p(|r|) = 6.88 \left(\frac{|r|}{r_0} \right)^{5/3} \tag{5.10}$$

The effects of the turbulent atmosphere are usually simulated using thin phase screens which perturb the phase of a propagating wave-front in accordance with Kolmogorov theory . Although there are several techniques for generating phase screens, the most popular methods are based on filtering white Gaussian noise in the spectral domain and then transforming to the spatial domain using the Fast Fourier Transform (FFT). Hence, a discrete phase screen can be generated according to

$$\varphi(x, y) = \sum_{k_x} \sum_{k_y} h(k_x, k_y) \sqrt{\Phi_p(k_x, k_y)} e^{i(k_x x + k_y y)} \Delta k_x \Delta k_y \qquad (5.11)$$

The discrete spatial domain points x and y are given by $x = m\Delta x$ and $y = n\Delta y$, where Δx and Δy are the desired sample intervals, and m and n are integer indices. The discrete wave number domain sample points k_x and k_y are given by $k_x = m'\Delta k_x$ and $k_y = n'\Delta k_y$, where Δk_x and Δk_y are the sample intervals, and m' and n' are integer indices. The discrete white noise process is given by

$$h(k_x, k_y) = \frac{g(m'n')}{\sqrt{\Delta k_x \Delta k_y}} \qquad (5.12)$$

where $g(m', n')$ is a discrete Gaussian noise process defined as:

$$g(m', n') = g_{real}(m', n') + i g_{imag}(m', n'), \qquad (5.13)$$

and $g_{real}(m', n')$ and $g_{imag}(m', n')$ are zero-mean Gaussian deviates with a standard deviation of $1/\sqrt{2}$. Because $h(m', n')$ is Hermitian, $g(m', n') = g^*(-n', -m')$. The scaling by $\sqrt{\Delta k_x \Delta k_y}$ is required to make the correlation of the discrete noise process approximately a continuous 2D spatial delta function. To implement (5.11) with an FFT, we convert from wave number space to the spatial frequency domain ($k = 2\pi f$). Suppose that we define the x and y sizes of the screen by G_x and G_y, each with N_x and N_y points, respectively (N_x and N_y are both powers of two, as required by the FFT). The corresponding sample intervals are $\Delta x = G_x / N_x$ and $\Delta y = G_y / N_y$. The spatial frequencies are defined by $f_x = m'\Delta f_x$ and $f_y = n'\Delta f_y$, where $\Delta f_x = 1/G_x$ and $\Delta f_y = 1/G_y$. Substituting the discrete variables in (5.11) we obtain the following for discrete phase screen

$$\varphi(m, n) = \sum_{m'=-N_x/2}^{N_x/2} \sum_{n'=-N_y/2}^{N_y/2} h(m', n') f(m', n') \exp\left(i2\pi\left(\frac{m'm}{N_x} + \frac{n'n}{N_y}\right)\right) \qquad (5.14)$$

where

$$f(m', n') = \frac{2\pi}{\sqrt{G_x G_y}} \sqrt{0.00058} r_0^{-5/6} \left(f_x^2 + f_y^2\right)^{-11/12} \qquad (5.15)$$

is the turbulence spatial filter and $h(m', n')$ is the white noise process. We see that the phase screen $\varphi(m, n)$ is simply the inverse FFT of the product of filter function and a white noise realization.

Although simple to implement, the FFT-based method do not span all the spectral frequency ranges corresponding to inner and outer scales of turbulence eddies. For the FFT-based phase screen to include the entire spectral range, a huge grid with extremely large sampling rate is required. Ideally, the phase screen must be as large as the outer scale of turbulence L_0, with sample-spacing as small as l_0. However, L_0 is several orders of

magnitude larger than l_0, and hence a large number of samples is required for this method to work properly. If the number of samples is not sufficient, due to abrupt change of power spectrum in the vicinity of origin, low frequencies can not be properly represented in the phase screen. One remedy to this problem is the non-uniform sampling of the frequency spectrum [40][41][42]. New techniques, such as sub-harmonics method, or random mid-point displacement algorithm have been proposed to overcome these limitations [40][41][42][43][44]. Another method, which is popular for circular apertures and is based on orthogonal expansion of phase screen using Zernike polynomials is also widely used in the literature [6][45]. Next, phase screen generation methods are described and compared to one another based on their compliance with the Kolmogorov spectrum and structure function, as well as their computational complexity.

5.3 Sub-Harmonics Method

Kolmogorov spectrum changes very rapidly near the origin and hence a single sample at the origin does not properly model the Fourier transform. To solve this problem, sub-harmonic method replaces the sample at the origin with nine samples at (-1/3, -1/3), (-1/3, 0), (-1/3, 1/3), (0, -1/3), etc. However, the sub-patches corresponding to these nine samples have 1/9 of the area of original path size. Sample points are placed in eight outer sub-patches, and the process can be continued by replacing the remaining central sub-patch with the secondary sub-patches. Hence, we can generate several low-frequency phase screens and add them to the high frequency grid. Fig. 5.6 shows the relationship between the high and low frequency spectra.
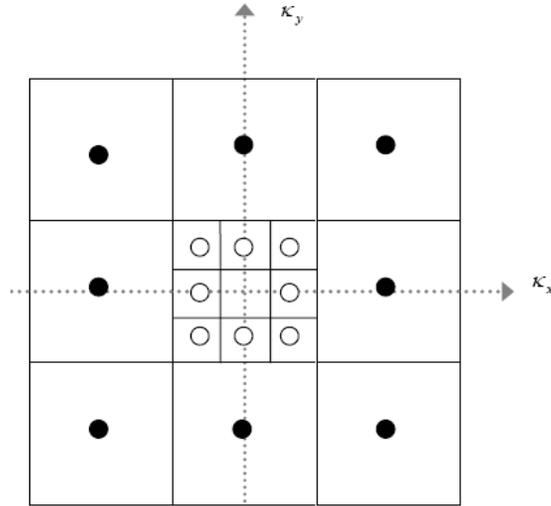
Fig. 5.6: The relationship between the high and low frequency spectra in the sub-harmonic method [41].

Assuming N_p sub-harmonics level, the low frequency screen is expressed as:

$$\phi_{LF}(m,n) = \sum_{p=1}^{Np} \sum_{m'=-1}^{1} \sum_{n'=-1}^{1} h(m',n') f(m',n') \exp\left(i2\pi 3^{-p}\left(\frac{m'm}{N_x} + \frac{n'n}{Ny} \right)\right) \qquad (5.16)$$

where

$$f(m',n') = \frac{2\pi 3^{-p}}{\sqrt{G_x G_y}} \sqrt{0.00058} r_0^{-5/6} (f_x^2 + f_y^2)^{-11/12} \qquad (5.17)$$

The values of modified screen are obtained by the summation of the high and low frequency grids, i.e.

$$\varphi_{\mathrm{mod}}(m,n) = \varphi(m,n) + \varphi_{LF}(m,n) \qquad (5.18)$$

The sub-harmonics method has different versions depending on the number of sub-patches at each stage, but the idea is basically the same. Fig. 5.7 shows an FFT-based phase screen, while Fig. 5.8 shows the same phase screen after addition of 12 sub-harmonics to represent the lower frequencies. One can easily observe the tilt added to phase screen in Fig. 5.8.
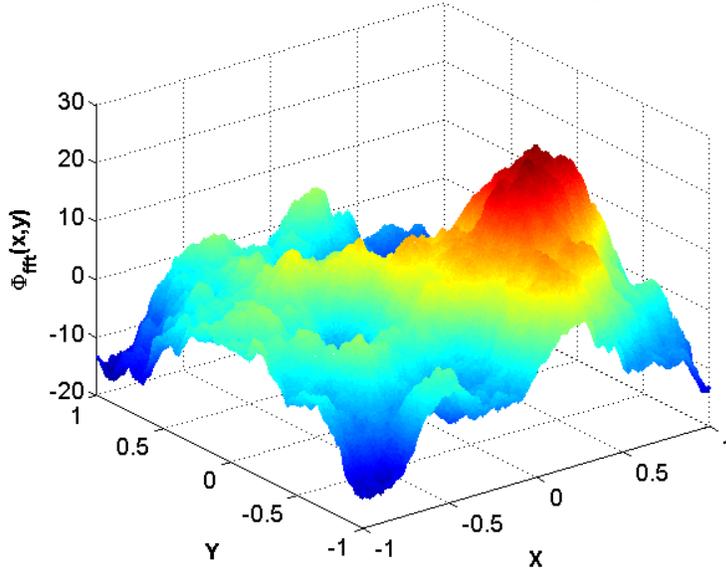
Fig. 5.7: FFT-based phase screen.
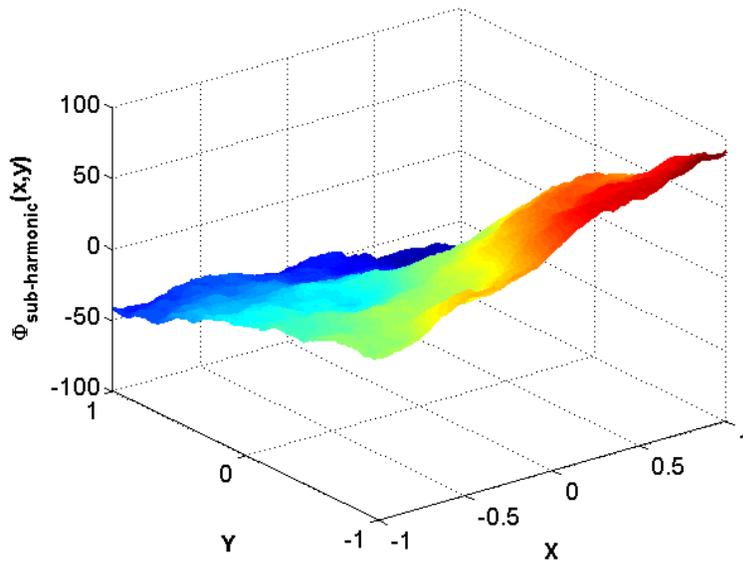
Fig. 5.8: FFT-based phase screen after addition of sub-harmonics.

Fig. 5.9 shows how by increasing the number of sub-harmonics, structure function of phase screens becomes closer to the structure function given by Kolmogorov theory.
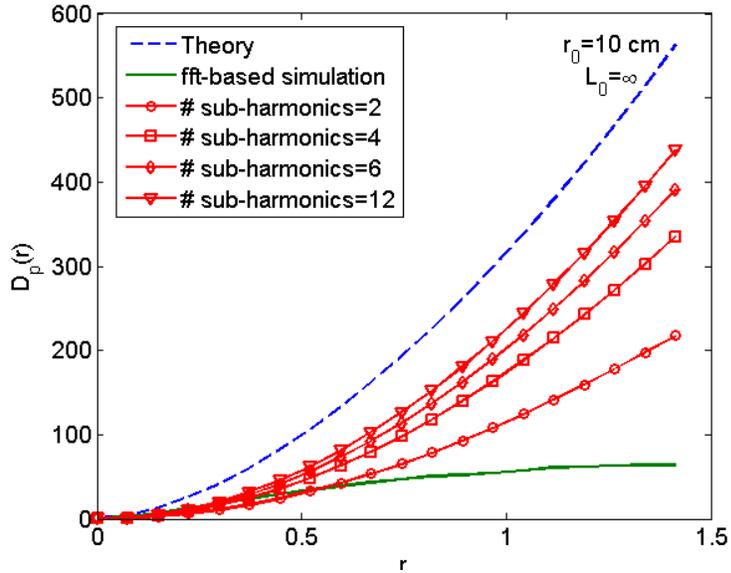
Fig. 5.9: Comparison of the sub-harmonic method structure function with theory.

5.4   Random Mid-Point Displacement Algorithm

The sub-harmonics method presented in the previous section modifies the original phase screen by adding the contributions from low frequencies. However, the limited number of sub-harmonics adds just a few low-frequency oscillations in limited directions. We can overcome this problem by exploiting the fractal nature of the phase screens. The latter implies that phase screens are self-similar in $l_0 < \rho < L_0$, regardless of the scale they are viewed in. In other words, unlike the conventional curves that look smoother when viewed in small scale and become rougher when viewed in large scale, the phase screens look the same in the above-mentioned range. Hence, we can generate an initial phase screen on a rectangular grid of a diameter in the order of L_0, but with a small number of samples. Then by interpolation of the central part of the initial phase screen, a new phase screen with finer sampling is generated. This process is continued, until the diameter of the phase screen becomes as small as the aperture diameter. To retain the random nature of the phase screen, in each iteration, random displacements should be added to the interpolated points [40][44].
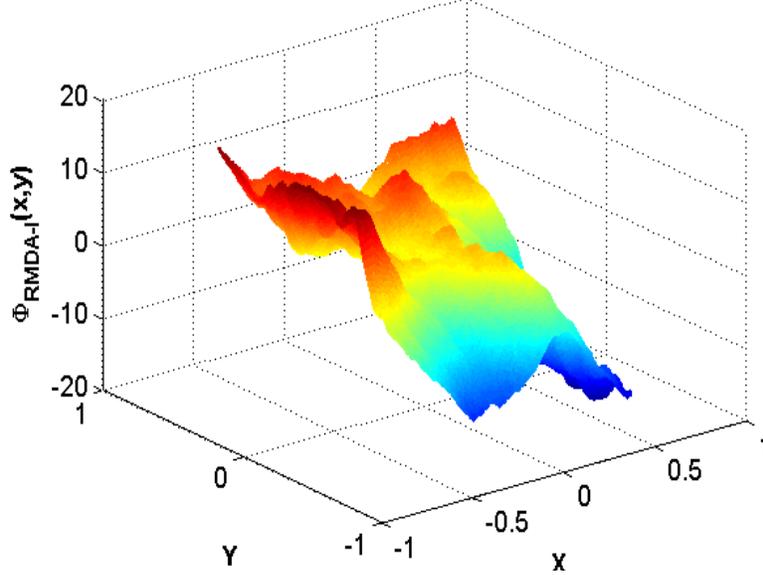
Fig. 5.10: Sample phase screen generated using RMDA-I.

Hence, this class of phase screen generation methods uses some version of Random Mid-point Displacement Algorithm (RMDA). This algorithm has been widely used in the literature for generating fractals [46]. There are many variations of this algorithm in the literature; however, in this chapter we describe only two major versions of this algorithm.

The first variation of RMDA, called RMDA-I from now on, starts with generation of a large FFT-based phase screen. Then by interpolating and adding another FFT-based phase screen as a random displacement, a smaller grid is formed. Repeating this process for several steps leads to a phase screen of desired structure function. Table 5.1 summarizes the steps representing RMDA-I for generating a phase screen [44]. Fig. 5.10 shows a sample phase screen generated using RMDA-I.

The second version of RMDA, called RMDA-II in this chapter, also starts with making a rough approximation to the fractal surface and then performing successive refinements in smaller localized portions. As described in [40] and [46], in order to simulate Kolmogorov turbulent layer, we start with four initial samples α, β, γ, and δ, generated randomly according to Kolmogorov structure function, as shown in Fig.5.11. In order to reflect the correlation between the starting samples, α, β, γ, and δ are generated from six random variables as:

$$
\begin{aligned}
\alpha &= R_\alpha^c + 0.5 R_{\alpha\delta}^d \\
\beta &= R_\beta^c + 0.5 R_{\beta\gamma}^d \\
\gamma &= R_\gamma^c - 0.5 R_{\beta\gamma}^d \\
\delta &= R_\delta^c - 0.5 R_{\alpha\delta}^d
\end{aligned}
\tag{5-19}
$$

Where R_α^c, R_β^c, R_γ^c, and R_δ^c are independent zero-mean Gaussian random variables of variance σ_c^2, and $R_{\alpha\delta}^d$ and $R_{\beta\gamma}^d$ are independent zero-mean Gaussian random variables of variance σ_d^2, added to represent correlation between the samples. From structure function of (5.10), we know that:

54

$$\left\langle (\alpha - \beta)^2 \right\rangle = 2\sigma_c^2 + \frac{\sigma_d^2}{2} = 6.88\left(\frac{D}{r_0}\right)^{5/3}$$

$$\left\langle (\alpha - \delta)^2 \right\rangle = 2\sigma_c^2 + \sigma_d^2 = 6.88\left(\frac{\sqrt{2}D}{r_0}\right)^{5/3}$$

(5.20)

Table 5.1: Random Mid-point Displacement Algorithm, Version I.

- **Set** L_0, D, r_0, and N. Usually L_0 is set to $2^n D$, where n is an integer.
- **Set** part=1/2, (the part to be extracted from the center of the phase screen)
- **Create** an initial phase screen covering the outer scale area:

    $x_{max} = L_0/2, \Delta x = L_0/N,$

    $f_{max} = 1/(2\Delta x), \Delta f = 1/L_0.$

    Do:

    Extract the central part of the previously created phase screen.
    Interpolate the extracted screen to generate an $N \times N$ grid. The interpolation can be done by cubic interpolation.
    Calculate new parameters:

    $x_{max} = x_{max} \times part, \Delta x = \Delta x \times part,$

    $f_{min} = f_{max}, \Delta f = 1/(2x_{max}), f_{max} = 1/(2\Delta x)$

    Create a new phase screen with nonzero spectral component in the interval $[f_{min}, f_{max}]$.
    Add both screens (new and interpolate)

    Repeat until $2x_{max} = D$.

Hence, one can show that:

$$\sigma_c^2 = 0.7506\left(\frac{D}{r_0}\right)^{5/3}$$

$$\sigma_d^2 = 10.7575\left(\frac{D}{r_0}\right)^{5/3}$$

(5.21)

A central sample m is formed by interpolating the corner samples α, β, γ, and δ plus a random displacement, ε as:

$$m = \frac{\alpha + \beta + \gamma + \delta}{4} + \varepsilon \qquad (5.22)$$

The variance of ε is chosen with respect to the structure function. Knowing that

$$\left\langle (\alpha - m)^2 \right\rangle = \frac{3\sigma_c^2}{4} + \frac{\sigma_d^2}{4} + \sigma_\varepsilon^2 = 6.88 \left(\frac{D/\sqrt{2}}{r_0} \right)^{5/3}, \qquad (5.23)$$

variance of ε turns out to be $\sigma_\varepsilon^2 = 0.6091(D/r_0)^{5/3}$. Next, edge samples are generated in mid-way of corner samples, using interpolation and random displacement. This process is continued to generate a large grid, as shown in Fig. 5.12.
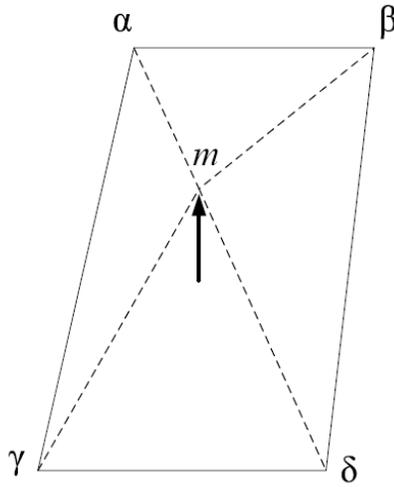
Fig.5.11: Four starting samples in RMDA-II and the central sample generated by interpolation and random displacement [40].

Fig. 5.13 shows a sample phase screen generated using RMDA-II. More over Fig. 5.14 illustrates the ensemble radial average of the phase structure function of 500 sample phase screens generated using RMDA-I and RMDA-II. One can see that for RMDA-I, by increasing the number of interpolations, structure function gets closer to curve obtained from theory.

Fig. 5.12: The interpolation sequence of RMDA-II (a) for initial samples are used to generate the central sample, (b) the corner samples are used to generate four interpolated edge samples (shown by filled-in dots) to produce a 3×3 grid. (c) and (d) show steps needed to generate a 5×5 grid from a 3×3 one [40].

Fig. 5.13: Sample phase screen generated using RMDA-II.
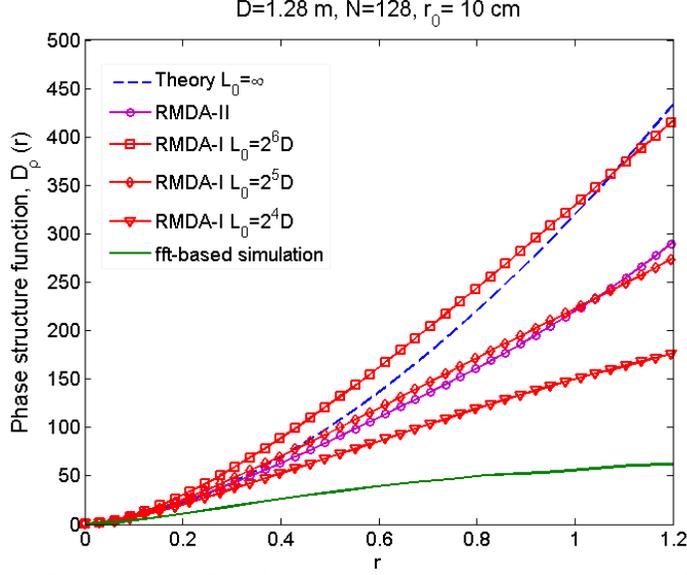
Fig. 5.14: Comparison RMDA method structure function with theory

5.5   *Phase Screen Generation Using Zernike Polynomials*

Another widely used phase screen generation method is based on expansion of the phase front into a linear combination of a set of ortho-normal functions over the aperture [45][47][48]. Zernike polynomials are a set of polynomials defined on a unit circle and are represented in polar coordinate as a product of radial polynomials and angular functions. The phase perturbation of an optical wave-front $\psi(\rho,\varphi)$ can be expanded as:

$$\psi(\rho,\varphi) \approx \sum_{i=1}^{N} a_i Z_i(\rho,\varphi) \qquad (5.24)$$

where $Z_i(\rho,\varphi)$ is the ith Zernike polynomials. The approximately equal sign in (5.24) is due to the fact that the phase screen expansion is obtained only up to *N* modes, rather than an infinite sum. Since the Zernike polynomials are ortho-normal only on the unit circle in two dimensions, the radius of the aperture on which the phase perturbation is generated must be normalized to one as $\rho = r/R$, where *R* is the radius of the aperture.

Mathematically, the ortho-normality condition that is satisfied by the Zernike polynomials can be expressed as:

$$\int_{\rho=0}^{1} \int_{\theta=-\pi}^{\pi} \rho d\rho d\theta W(\rho) Z_i(\rho,\theta) Z_j(\rho,\theta) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \qquad (5.25)$$

That is, the polynomials are orthogonal only over the unit circle. The function $W(\rho)$ is the aperture weighting function or pupil function defined so that $W(\rho)=0$ outside the telescope pupil, and furthermore it is normalized so that $\int_{r=0}^{1} \int_{\theta=-\pi}^{\pi} \rho d\rho d\theta W(\rho)=1$. The coefficients $\{a_i\}$ are given by the projections of the phase perturbation $\psi(\rho,\varphi)$ on the basis polynomials $\{Z_i(\rho,\theta)\}$, i.e.,

58

$$a_i = \int_{(\rho,\theta)} \rho d\rho d\theta W(\rho) \psi(R\rho,\theta) Z_i(\rho,\theta) \tag{5.26}$$

A convenient scheme of ordering the polynomials proposed by Noll [45], and adopted by most literature, is followed here. As such, the ith member of the Zernike polynomial set, $Z_i(\rho,\theta)$ can be expressed as

$$Z(\rho,\theta) = \begin{cases} \sqrt{n+1} R_n^m(\rho)\cos(m\theta), \text{ for } i = \text{even}, m \neq 0 \\ \sqrt{n+1} R_n^m(\rho)\sin(m\theta), \text{ for } i = \text{odd}, m \neq 0 \\ R_n^0(\rho), \text{ for } m = 0 \end{cases} \tag{5.27}$$

Radial functions $R_n^m(\rho)$ are defined by,

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s![(n+m)/2-s]![(n-m)/2-s]!} \rho^{n-2s} \tag{5.28}$$

Azimuthal and radial orders, m and n, are non-negative integers so that $m \leq n$ and $(n-m) = \text{even}$. Some low order Zernike modes are demonstrated in Fig. 5.15.

Turbulence-induced phase perturbation, $\psi(R\rho,\theta)$, can be expanded, using Zernike polynomials as stated in (5.24). Given central limit theorem and Gaussian nature of phase aberration, random variables a_i are Gaussian distributed. If the phase perturbation $\psi(R\rho,\theta)$ is zero-mean, then the mean value of $a_i, \forall i$ is:

$$\begin{aligned} E\{a_i\} &= E\left\{ \int_{(\rho,\theta)} \rho d\rho d\theta W(\rho) \psi(R\rho,\theta) Z_i(\rho,\theta) \right\} \\ &= \int_{(\rho,\theta)} \rho d\rho d\theta W(\rho) E\{\psi(R\rho,\theta)\} Z_i(\rho,\theta) = 0 \end{aligned} \tag{5.29}$$

Similarly, the covariance values of the expansion coefficients can be found from

$$\begin{aligned} E\{a_i a_j\} &= E\left\{ \int_{(\rho,\theta)} \rho d\rho d\theta W(\rho) \psi(R\rho,\theta) Z_i(\rho,\theta) \int_{(\rho',\theta')} \rho' d\rho' d\theta' W(\rho') \psi(R\rho',\theta') Z_j(\rho',\theta') \right\} \\ &= \int_{(\rho,\theta)} \rho d\rho d\theta \int_{(\rho',\theta')} \rho' d\rho' d\theta' W(\rho) W(\rho') Z_i(\rho,\theta) Z_j(\rho',\theta') E\{\psi(R\rho,\theta) \psi(R\rho',\theta')\} \end{aligned} \tag{5.30}$$

Replacing phase covariance function in (5.30) and rewriting it in Fourier space we have [45]:

$$\langle a_i^* a_j \rangle = \iint dk dk' Q_i^*(k) \Phi(k/R, k'/R) Q_i(k') \tag{5.31}$$

where

$$\Phi(k/R, k'/R) = 0.023(R/r_0)^{5/3} k^{-11/3} \delta(k-k') \tag{5.32}$$

and $Q_i^*(k)$ and $Q_j(k')$ are Fourier transforms of Zernike modes. Hence, covariance values of the expansion coefficients turns out to be:

$$\overline{a_i a_j} = 0.0072 \left(\frac{D}{r_0}\right)^{5/3} (-1)^{(n_i+n_j-2m_i)/2} [(n_i+1)(n_j+1)]^{1/2} \pi^{8/3} \delta_{m_i m_j}$$

$$\times \frac{\Gamma(14/3)\Gamma[(n_i+n_j-5/3)/2]}{\Gamma[(n_i-n_j+17/3)/2]\Gamma[(n_j-n_i+17/3)/2]\Gamma[(n_i+n_j+23/3)/2]} \tag{5.33}$$

for i - j=even and for i - j=odd

$$\overline{a_i a_j} = 0 \tag{5.34}$$

where, m_i and n_i refer to the azimuthal and radial orders associated with the ith Zernike polynomial, and δ_{ij} is the Kronecker delta function. Table 5.2 lists the coefficients of (5.33) and (5.34) [6].
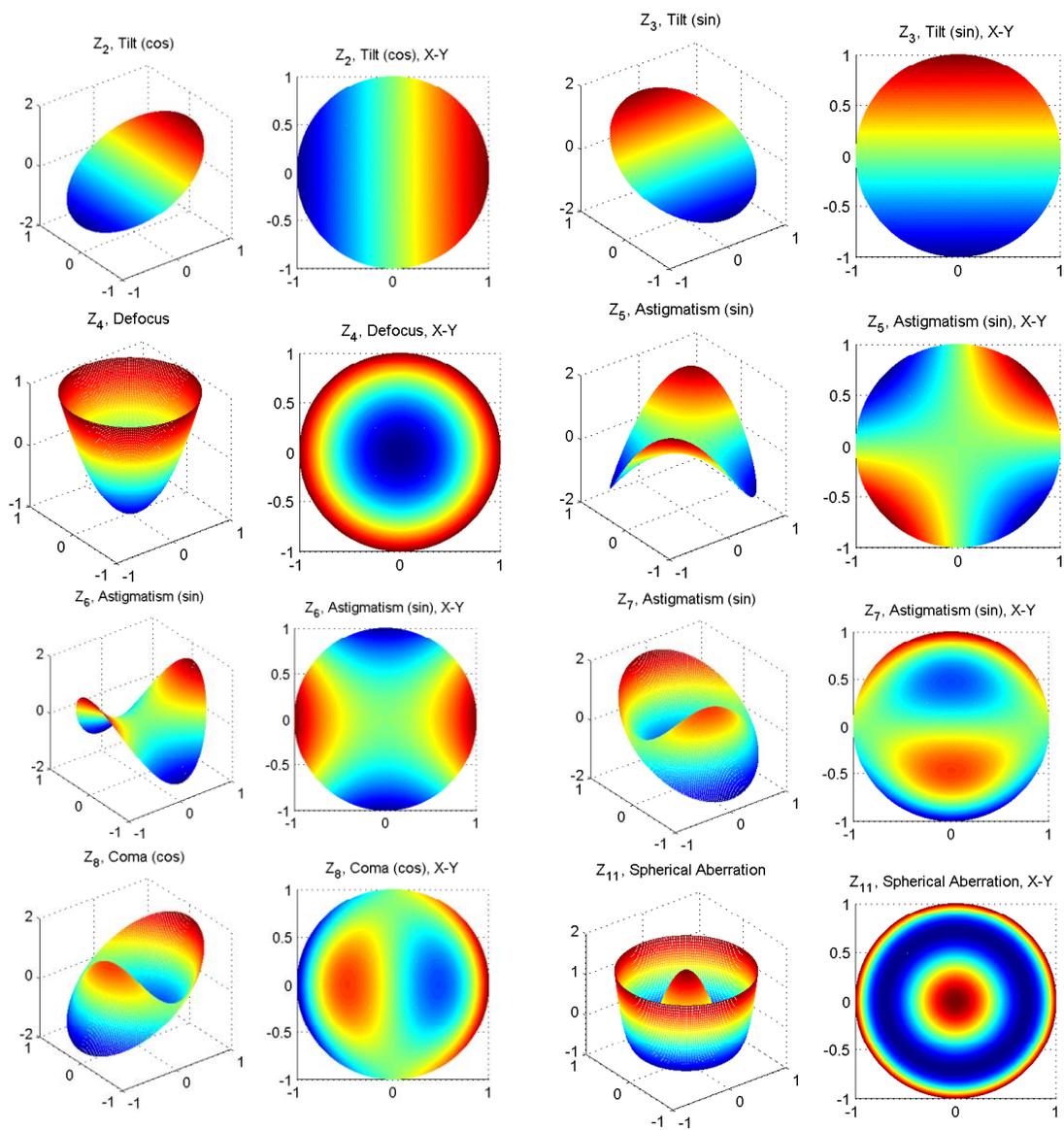
Fig. 5.15: Zernike modes for tilt, defocus, astigmatism, coma, and spherical aberration.

Table 5.2: $\overline{a_i a_j}(D/r_0)^{-5/3}$ for the first 10 Zernike polynomials. Rows correspond to i and columns correspond to j.

i / j	2	3	4	5	6	7	8q	9	10
2	0.448	0	0	0	0	0	-0.0141	0	0
3	0	0.448	0	0	0	-0.0141	0	0	0
4	0	0	0.0232	0	0	0	0	0	0
5	0	0	0	0.0232	0	0	0	0	0
6	0	0	0	0	0.0232	0	0	0	0
7	0	-0.0141	0	0	0	0.00618	0	0	0
8	-0.0141	0	0	0	0	0	0.00618	0	0
9	0	0	0	0	0	0	0	0.00618	0
10	0	0	0	0	0	0	0	0	0.00618

Fig. 5.16 shows a sample phase screen generated using modal expansion. Moreover Fig. 5.17 demonstrates the radially averaged phase structure function of an ensemble of 500 phase screen generated using Zernike polynomials.
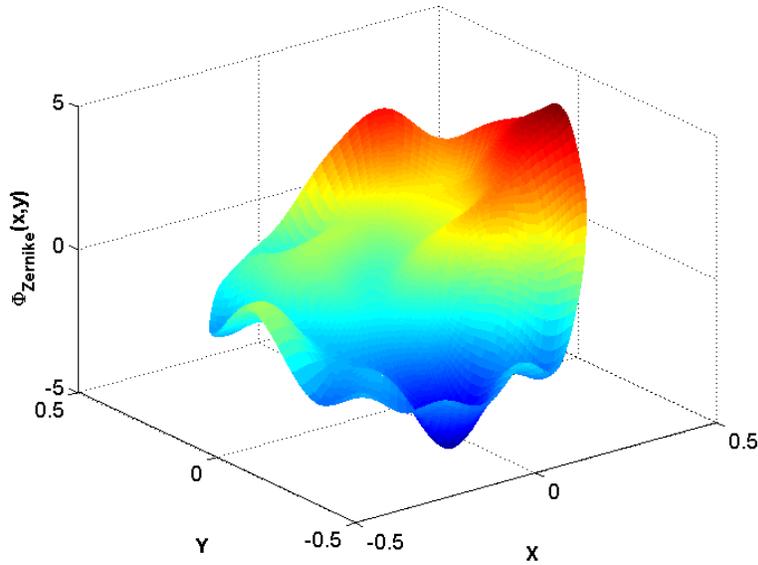
Fig. 5.16: Sample phase screen generated using Zernike polynomials, $D=1.28$, $r_0=10$ cm.
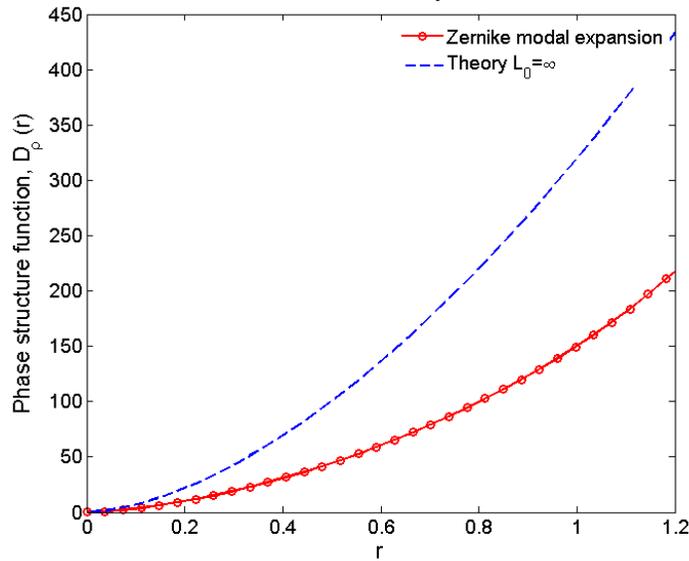
Fig. 5.17: Radially averaged structure function for an ensemble of 200 phase screens, using modal expansion, *D=1.28, r₀=10 cm*.

Table 5.2 show a small amount of correlation between coefficients of Zernike polynomials, used to generate phase screens, suggesting that there may be a more efficient set of orthonormal basis function with uncorrelated coefficients. In fact Fried has solved this problem in [49], where Karhunen-Loeve integral equation is solved. It is also possible to obtain this expansion from linear combinations of Zernike polynomials, providing basis functions which are optimal in the sense that the coefficients of the expansion are uncorrelated. To this end, the eigen-decomposition of the covariance matrix of the Zernike coefficients is calculated as:

$$\Gamma_a = U \Lambda U^T \tag{5.35}$$

where U is the matrix with columns corresponding to the eigenvectors, and Λ is a diagonal matrix containing eigen values of the Hermitian covariance matrix.

The ortho-normal polynomials (with the first order polynomial corresponding to piston removed) are obtained by forming the sum of the Zernike polynomials weighted by each element of the columns of U, as:

$$E_j(\rho,\theta) = \sum_{i=2}^{N} U_{ij} Z_i(\rho,\theta) \tag{5.36}$$

Thus, phase perturbation can be expanded as:

$$\phi(R\rho,\theta) = \sum_{i=2}^{N} b_i E_i(\rho,\theta) \tag{5.37}$$

where the b_i's are uncorrelated Gaussian random variables with variances given by the diagonal elements of Λ. Some low order orthonormal polynomials using eigendecomposition of covariance of Zernike coefficients upto 300 terms is illustrated in Fig. 5.18.
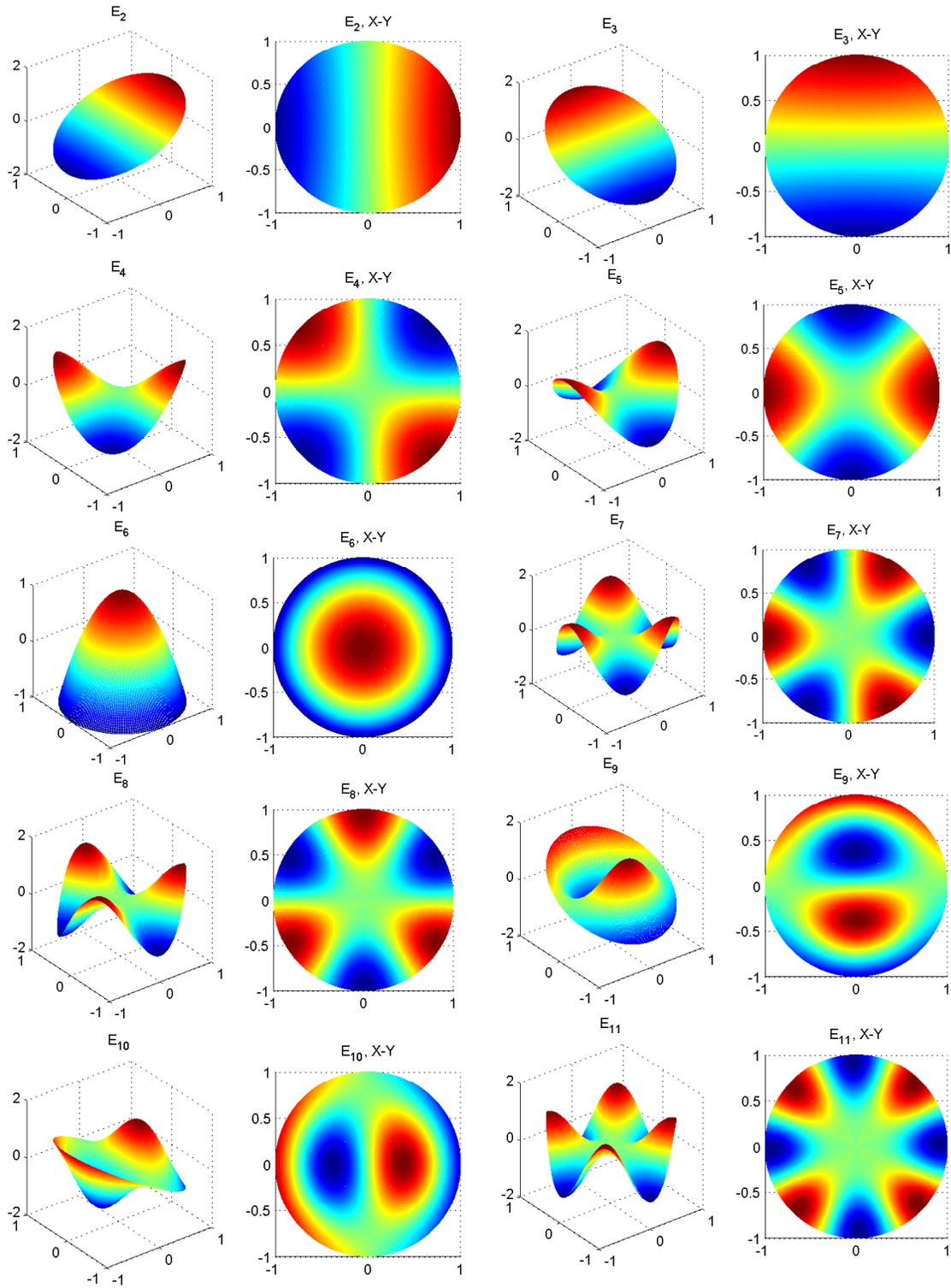
Fig. 5.18: Some Low Order Ortho-normal polynomials using eigen decomposition of covariance matrix of Zernike coefficients up to 300 terms.

Given the order N of expansion and the corresponding covariance matrix of the Zernike coefficients for a specified D/r_0 ratio, Gaussian-distributed random realizations of the

coefficients $a_i, i \neq 1$ can be generated for the piston-free phase screen. We can obtain the Cholesky decomposition of the covariance matrix Γ_a, F_a as,

$$\Gamma_a = F_a^T F_a \tag{5.38}$$

Then the Gaussian-distributed vector $\mathbf{a}$ with the required covariance of (5.33) can be obtained as,

$$\mathbf{a} = F_a^T \mathbf{g} \tag{5.39}$$

where $\mathbf{g}$ is a $(N-1) \times 1$ vector with uncorrelated random variables each having a Gaussian distribution with zero-mean and unit-variance. The one less term is due to removal of piston. Similarly, the KL ortho-normal polynomial expansion can also be used to obtain realizations of phase screens after generating a coefficient vector $\mathbf{b}$ as follows,

$$\mathbf{a} = F_a^T \mathbf{g} \tag{5.40}$$
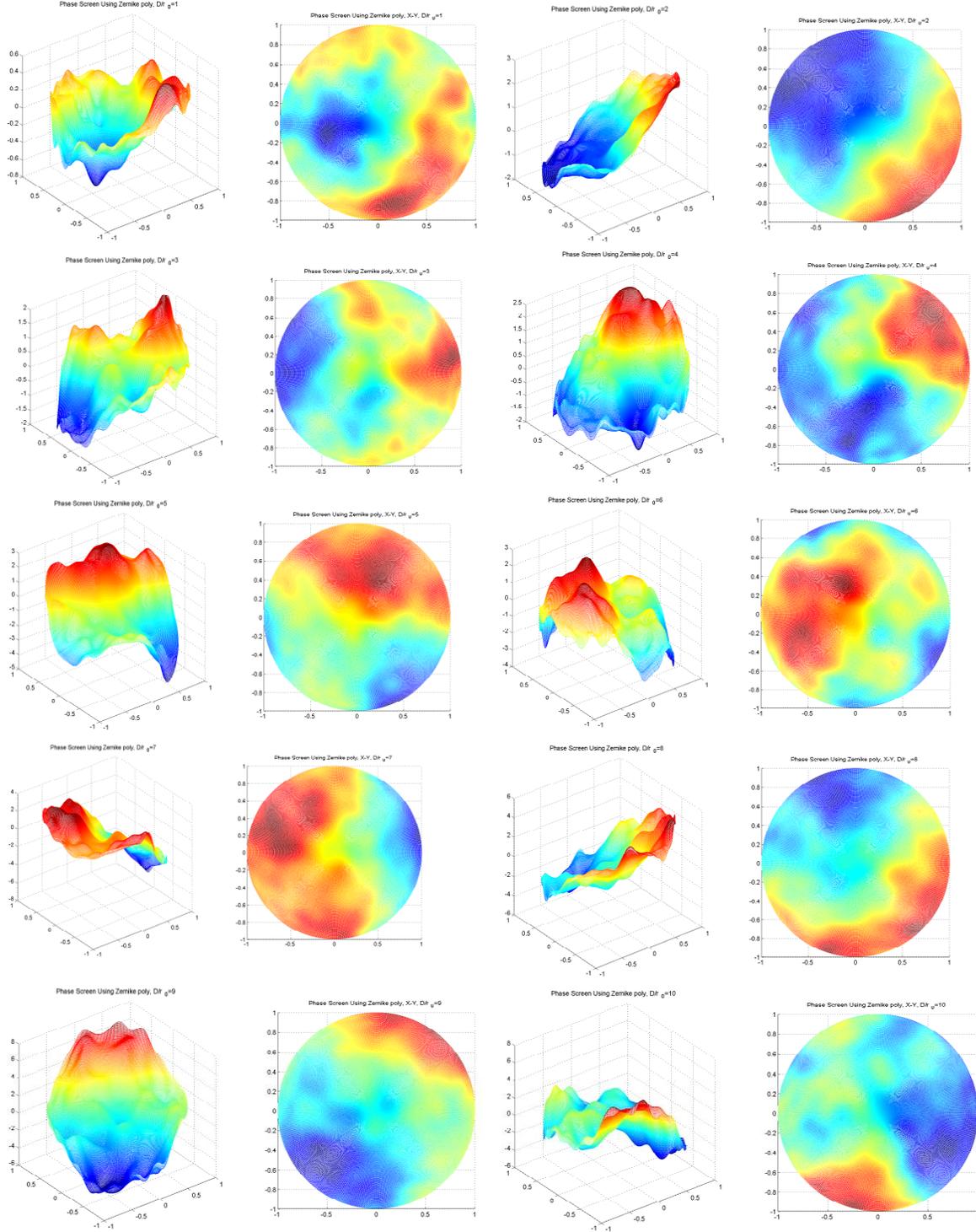
Fig. 5.19: Phase Screen Realizations for different D/r_0 ratios.

where Λ is the diagonal matrix from (5.35), and **g**, as before is a $(N-1) \times 1$ vector with uncorrelated random variables each having a Gaussian distribution with zero-mean and unit-variance. Then, a random realization of the phase screen is obtained similar to (5.37).

The phase screens obtained by both expansions are almost equivalent, apart from some mixing of higher order Zernike modes in the orthonormal expansion. Some random realizations for different D/r_0 ratios are demonstrated in Fig. 5.19.

5.6 Comparison of Phase Screen Generation Techniques

Different techniques for generating phase screens are discussed in this chapter. Each of these techniques is suitable for a different optical system and channel setup. For example, Zernike modal expansion method is useful for circular apertures and structure functions of Sub-Harmonics and RMDA methods approaches to theoretical curves, by increasing the number of sub-harmonics and interpolations. However, it is worthwhile to compare these techniques from computational complexity point of view. To this end, time required for generating a sample 128×128 phase screen, over an aperture of 1.28 m diameter, for an accumulated Fried's parameter of 10 cm, is measured. Table 5.3 lists the running times.

Table 5.3: Running Time for Different Phase screen Generation Method.

Method	Sub-Harmonics (6 sub-harmonics)	RMDA-I (6 interpolations)	RMDA-II	Zernike expansion (300 modes)
Time (seconds)	2.45	0.36	0.05	1.30

As one can see from Table 5.3, RMDA-I and RMDA-II, and specifically RMDA-II are much faster than other phase screen generation techniques. In other words, if a simulation needs a large ensemble of phase screens, using RMDA-II method is much more efficient. Note that the codes used to generate these phase functions are not professionally optimized to minimize the execution time. In the next chapters, phase screen generation methods are widely used to simulate atmospheric turbulence and measure performance of optical imaging systems' in different atmospheric conditions.

5.7 Conclusions

In this chapter, atmospheric turbulence models are described and the impact of refractive index variations across the optical path on the received wave-front is investigated. It is shown that turbulence can be modeled by one or several thin phase screens which are generated according to Kolmogorov spectrum and different methods of generating such phase screens are discussed and compared based on accuracy and computational complexity. Having investigated both the MCRT algorithm and phase-screen generation methods, now we have a comprehensive channel model that can be exploited in design and evaluation of optical communications and imaging systems. In the next chapters, atmospheric active imaging systems are discussed using this framework.

Chapter 6
Spatially Multiplexed Optical MIMO Imaging System in Cloudy Turbulent Atmosphere

6.1   Introduction

Active optical imaging systems offer higher resolution and faster area search rate compared to radio frequency (RF) counterparts due to their considerably smaller wavelengths. Furthermore, since optical and infra-red wavelengths are closer to visible range, the image quality is similar to that of visible, and hence easier to interpret by a human observer. However, obstruction by clouds and fog is, and will continue to be, the major limitation of electro-optical/infrared (EO/IR) active imaging systems. Moreover, variations of refractive index in different layers of atmosphere, due to thermal changes and wind, introduce fluctuations in amplitude and phase of optical waves. In imaging through atmosphere one should consider dispersive effects of multiple scatterings, turbulence-induced beam spreading, intensity fluctuations, and wave-front distortions. All these phenomena broaden and distort impulse response of the imaging system, known as the Point Spread Function (PSF).

In this chapter, a spatially multiplexed Multi-Input Multi-Output (MIMO) imaging system configuration is proposed which uses a photolithographic beam splitter for target illumination in a pixel-wised manner. This yields a uniformly distributed illumination over the target area. Furthermore, beam spreading is reduced by a factor of *1/N*, where *N* is the number of beamlets in each direction [50] (i.e., we have an *N×N* array of beamlets) leading to a noticeable improvement in the link margin. The reflected beamlets can be detected by individual photo-detectors on a receiving Focal Plane Array (FPA). Hence, the problem of restoring a distorted image is significantly simplified and furthermore, considerable multiplexing gain may be achieved.

To compensate for atmospheric phenomena such as clouds, fog, aerosols, and even turbulence, MIMO imaging system should be well equipped with image restoration modules. While time-gating is suggested to filter-out back-scattered photons and increase the contrast, Adaptive-Optics (AO) and Blind Deconvolution (BD) modules are exploited to enhance the image sharpness and resolution.

6.2   Atmospheric Channel Modeling

In active optical imaging systems, a laser beam is pointed towards a target and rays reflected back from the target are received and processed to retrieve an image. In the presence of scattering particles such as clouds and aerosols along the path of a pulsed laser, beams start to broaden in space and time and, hence, received power is attenuated. Furthermore, depending on the cloud optical depth, some photons are backscattered towards the receiver and contribute to a steady background noise. As a result, scattering media distort PSF of an imaging system due to pulse spreading and dispersion and also give rise to clutter due to backscattering. This clutter must be removed to increase the contrast or Signal-to-Noise Ratio (SNR). Turbulence on the other hand, induces fluctuations into the signal amplitude, which in turn causes fading of the received image. Furthermore, one should also consider the phase variations which distort the wave-front (Iso-phase plane) and PSF. This phenomenon is responsible for the speckle patterns formation.

We make a simplifying assumption that turbid and turbulent essences of atmosphere independently contribute to the PSF of an imaging system and hence can be modeled, separately. Later on, we show that given the proposed imaging system design, this is a valid assumption, since multiple scatterings can be modeled by a pure attenuation factor. While a Monte-Carlo Ray Tracing (MCRT) algorithm, developed at Pennsylvania State University, Center for Information and Communications Research (CICTR), is used to account for propagation through scattering media, turbulence is modeled by phase screens resulting from Rytov approximation and Kolmogorov theory of turbulence. Fig. 6.1 shows the geometry of an MCRT simulation for a mono-static imaging system.

Fig. 6.1: Geometry of MCRT for a mono-static imaging system.

The flowchart for the mono-static Monte Carlo Ray Tracing Imaging simulation is given in Table 6.1

Table 6.1: Flowchart of the mono-static Monte Carlo Ray Tracing for Imaging

```
                          ┌──────────────┐
                          │    Start     │
                          └──────┬───────┘
                   ┌─────────────┴──────────────┐
                   │          Define:           │
                   │ Global axis with origin at │
                   │     center of image.       │
                   │ Transmitter & receiver     │
                   │ local axes with origins &  │
                   │ tilts given in global      │
                   │ coordinates.               │
                   │ arrived_photons=0,         │
                   │ lost_photons=0             │
                   └─────────────┬──────────────┘
```

Define:
Global axis with origin at center of image.
Transmitter & receiver local axes with origins & tilts given in global coordinates.
arrived_photons=0, *lost_photons*=0

B

Enough photons reached receiver aperture?   YES →   Exit Function Output saved info

NO

Launch point:
Randomly pick photon launch point on transmitter plane.
Randomly pick initial launch angle & rotation.
count=1

First Scatter:
Randomly pick traveled distance, *d*, to next scatter site.

Calculate location of photon in global coordinates.

Crossed Image Plane?   NO →

YES

Reflected from image?   NO   YES   Absorbed at scatter site?

YES   NO

Set *image_reached* & *image_photon* flags to 1. Calculate reflection angle. Save current global location & scattering states.

lost_photons++ Set *end_trace* flag

Save current global location & scattering states

A

70

Table 6.1 (contd.): Flowchart of the mono-static Monte Carlo Ray Tracing for Imaging

```
                            ┌─────┐
                            │  A  │
                            └──┬──┘
                               │
                               ▼
                    ◇ end_trace=0? ◇ ──── NO ────▶ ⬠ B
                               │
                              YES
                               ▼
                    ┌──────────────────────┐
                    │       count++        │
                    │  Save previous location │
                    └──────────┬───────────┘
                   ┌───────────┴────────────┐
                   ▼                        ▼
          ◇ image_reached=1? ◇     ◇ image_reached=0? ◇
                   │                        │
                  YES                       ▼
                   ▼                      ⬠ C
      ┌────────────────────────┐
      │ Generate new scatter   │
      │ parameters, (d, θ, φ).  │
      │ (θ=reflection angle)   │
      │ Find new photon location w. r. │
      │   t. receiver axis.    │
      └───────────┬────────────┘
                  ▼
        ◇ Crossed Receiver plane? ◇ ──── NO ──────┐
                  │                                │
                 YES                               ▼
                  ▼                    ◇ Absorbed at scatter site? ◇
          ◇ Received? ◇ ── NO ──┐ YES ◀─────┘    │
                  │             │                 NO
                 YES            ▼                  ▼
```

Generate new scatter parameters, (d, θ, φ).
(θ=reflection angle)
Find new photon location w. r. t. receiver axis.

Set *end_trace* flag
arrived_photons++.
Save current & previous locations, total traveled distance, total number of scattering, & *image_photon* flag.

lost_photons++
Set *end_trace* flag

Save current global location & scattering states.
Reset *image_reached* flag to 0.

```
                           ┌─────┐
                           │  A  │
                           └─────┘
```

Table 6.1 (contd.): Flowchart of the mono-static Monte Carlo Ray Tracing for Imaging


```
                                    C

         Generate new scatter parameters, (d, θ, φ).
                   (θ from given PDF)
         Find new photon location w. r. t. receiver axis.
         Find new photon location w. r. t. global axis.

                    Crossed                NO
                 Receiver plane?

          YES

                                  Crossed Image          NO
         Received?                    Plane?
                        NO
          YES                          YES

   Set end_trace flag        NO                           Absorbed at      YES
   arrived_photons++                Reflected             scatter site?
   Save current & previous          from image?
   locations, total traveled
   distance, total number of         YES                  NO
   scattering, & image_photon
   flag                    Set image_reached &
                           image_photon flags to 1
                           Calculate reflection angle
                           Save current global location &
                           scattering states

                                                    Save current global
                                                    location & scattering
                                                    states

                  lost_photons++
                  Set end_trace flag

                                          lost_photons++
                                          Set end_trace flag

                        A
```

Since backscattered photons are of interest for clutter removal, they can be traced together with photons reflected from the target. Afterwards, the spatial and temporal distributions of both may be estimated in a post-processing stage. Our MCRT simulates a simplistic imaging

scenario, in which a 2-D mirror resembling the letter 'H' is placed at distance L from the pulsed laser source, i.e., transmitter. Assuming that both target and imaging system are inside the scattering medium, physical length of the channel, L, can be translated to its optical thickness, τ, as:

$$\tau = \beta_{ext} L = \frac{L}{D_{ave}},$$ (6.1)

where β_{ext} is the extinction coefficient of scattering medium and D_{ave} is the average distance between two successive scattering events. Hence, the round-trip optical thickness is equal to 2τ. We simulate this imaging scenario, assuming that laser beam path from transmitter to imaging plane and back to receiver goes through a scattering medium of the specified optical thickness. As a result of these scattering events, some of the energy is backscattered towards the transceiver. This backscattered energy appears as a clutter and hence, obscures the target. To remove the clutter space from the image space and improve the contrast, proper spatial and temporal processing have to be carried out. Fortunately, Back-scattered and image photons can be separated using time-gating due to the fact that back-scattered photons arrive much earlier than image photons, which have to travel the round trip path from the transceiver to the target and back. Using this time-gating approach, contrast is improved, significantly.

Using the setup shown in Fig. 6.1, an MCRT algorithm is developed, in which a large body of photons are traced until one million photons are either backscattered or reflected back from the target to the receiver. Fig. 6.2 shows spatial distribution of image and back-scattered photons at the receiving aperture for a channel of optical thickness 1 (2 round-trip). For a channel of this trivial optical thickness, number of image photons are much more than back-scattered ones. Fig. 6.3 shows the temporal distribution of image and back-scattered photons. It is clear that temporal filtering is capable of removing the clutter sub-space.
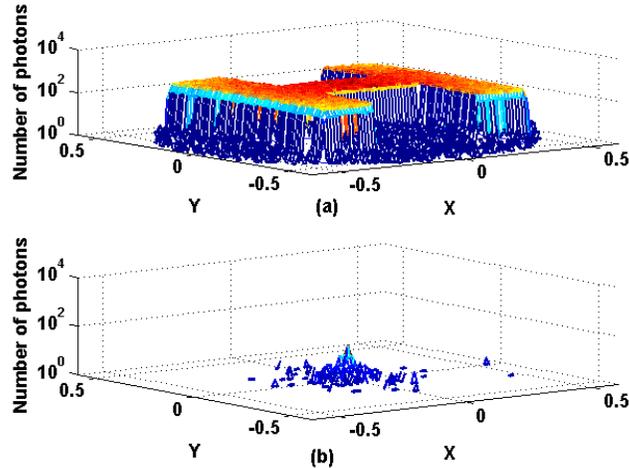
Fig. 6.2: Spatial distribution of (a) image photons and (b) backscattered photons at the receiver plane for a channel of optical thickness 1 (2 round-trip).
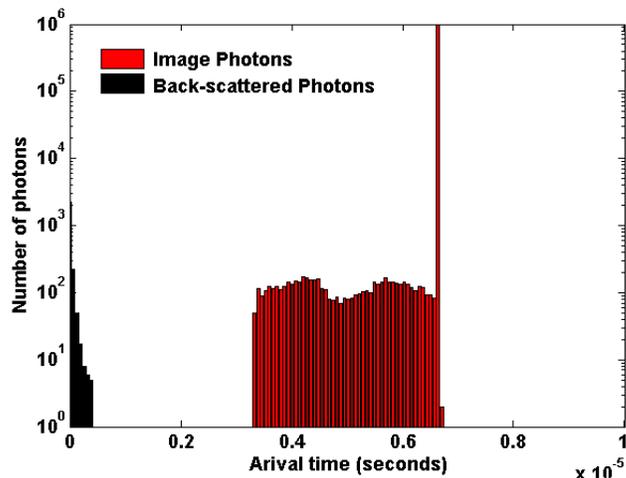
Fig. 6.3: Temporal distribution of image photons (red) and backscattered photons (black) at the receiver plane for a channel of optical thickness 1 (2 round-trip).

Fig. 6.4 and Fig. 6.5 show spatial and temporal distributions of photons for a channel of optical thickness 4 (8 round-trip). In this case, number of image photons is reduced, drastically. This is due to the fact that most information bearing photons are the so-called "ballistic" photons which are attenuated according to the Beer-Lambert law as:

$$I_{coh} = e^{-2\tau} I_0 \qquad (6.2)$$

where I_{coh} is the coherent part of the received radiance (Intensity) and I_0 is the transmitted pulse radiance (intensity). Photons that travel a close to Line-Of-Sight (LOS) path, i.e. "snake" photons, can also contribute to imaging and improve the link budget equation. However, special measures should be taken to employ these photons while maintaining the diffraction-limited resolution. It is worth noting that in [51] authors have used small-angle approximation to find the PSF of multiple-scattering medium. Back-scattering is not considered in their work; however, it was shown that for a receiver of large Field-Of-View (FOV), contrast and resolution are reduced, significantly. This is simply due to the fact that more diffuse photons are received and no spatial filtering is exploited.
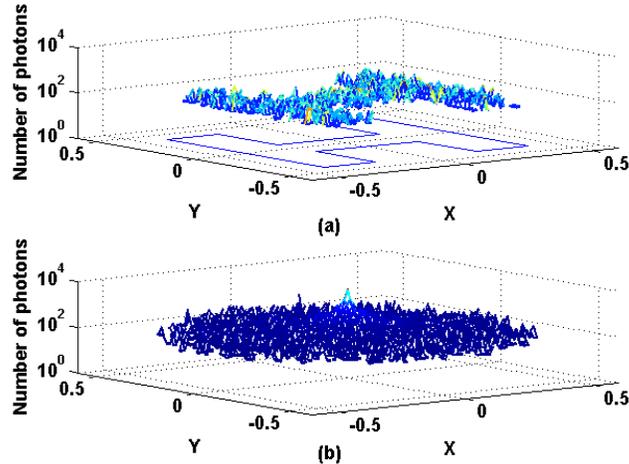
Fig. 6.4: Spatial distribution of (a) image photons and (b) backscattered photons at the receiver plane for a channel of optical thickness 1 (2 round-trip).
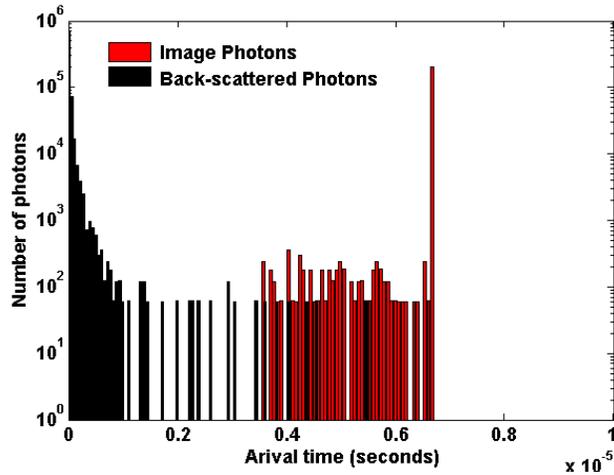
Fig. 6.5: Temporal distribution of image photons (red) and backscattered photons (black) at the receiver plane for a channel of optical thickness 1 (2 round-trip).

6.2.1   Bi-static Monte Carlo Imaging

A variant of the Monte Carlo Imaging technique discussed above is the bi-static case, where the transmitter and receiver are not collocated, rather they are separated by a certain amount. This type of imaging is convenient from the point of view of rejecting backscattered photons. The backscattered photons are projected more towards the transmitter and do not interfere with the photons reflected back from the target image. Fig. 6.6 and Fig. 6.7 indicate that the portion of backscatter photons reduce from 3% in the case of mono-static imaging with round trip optical thickness of 2, to 0.6%, which is a significant reduction. The transmitter and receiver, however, have to be tilted with appropriate angles to obtain proper alignment to capture the target image.
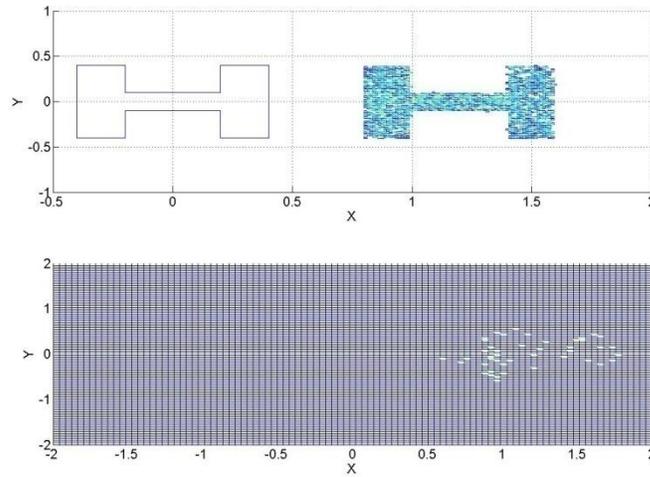
Fig. 6.6: Spatial distribution of (a) image photons and (b) backscattered photons at the receiver plane for a channel of optical thickness 1 (2 round-trip) for a bistatic system.
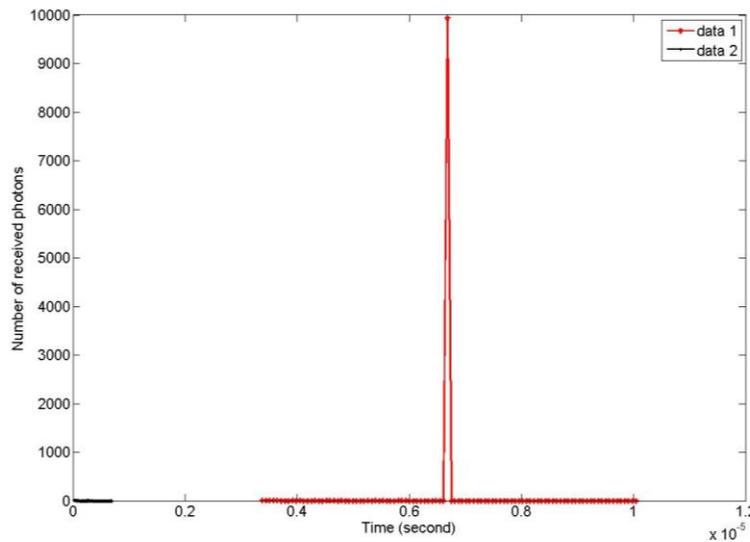
Fig. 6.7: Temporal distribution of (a) image photons and (b) backscattered photons at the receiver plane for a channel of optical thickness 1 (2 round-trip) for a bistatic system.

6.2.2   MIMO Monte Carlo Imaging

The major conclusion from the two different approaches for Monte Carlo imaging is that, by using a fast shutter and proper timing, we can mitigate the effect of the backscattered photons which are the primary source of noise in imaging through cloud and other scattering media. In this section, we present possible configurations for a MIMO imaging system through cloud and present the results for a 2×2 MIMO imaging system.

In a MIMO imaging system multiple beams are used to illuminate the target and multiple apertures are used to receive the reflections. The transmitters possibly need to be tilted so that they all point towards the target. Hence, the corresponding receivers must be placed where the

return rays are expected to arrive. There can be different possible geometric placements of transmitters and receivers. Also, due to the spatial confinement of the laser beams in clear weather conditions, parallel independent imaging system can be implemented by proper placement of transmitters and receivers. Even in the presence of an obscurant medium with a small optical thickness, such as a thin cloud, the line-of-sight component is always much larger than the diffuse component and hence the correlation between apertures is very small. One might expect high correlations between the received reflections in the case of optically thick obscurants; however, the power loss is so severe that even the intended receiver will not get much power under these circumstances. Fig. 6.8 shows an example of the possible geometry for the MIMO imaging system. The central transceiver is a mono-static imaging system, i.e. the same aperture is used for both sending the beam and receiving the reflections. The beams sent from other transmitters are supposed to be received by the reciprocal receivers.
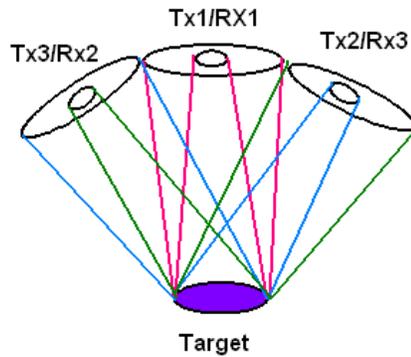
Fig. 6.8: An example of a possible geometry for a MIMO Imaging system.

A 2×2 MIMO imaging system composed of a mono-static and a bi-static SISO imaging systems is developed. The physical length of the channel is 1 km (2 km round trip) and the optical thickness is varied in each run of the simulation. Fig. 6.9 shows the received backscattered and image photons when the optical thickness of the medium is 2 (4 round trip). We see that the number of back-scattered photons is much smaller than the image photons. Furthermore, the backscattered photons are less for the bi-static imaging system, i.e. the second receiver. Another observation is that the density of the photons is not uniform all over the image. This phenomenon is more conspicuous when operating in a medium with a higher optical thickness. This is due to the randomness of the photons' trajectories. Due to this phenomenon, the signal-to-noise ratio (SNR) per pixel is not uniform all over the image and the quality in different portions of the image is not the same. By combining the images of different receivers, SNR per pixel improves and becomes more uniform all over the image. Both spatial and temporal characteristics of the received optical signals on different branches contribute to the possible image quality enhancement.

In Fig. 6.10, the arrival times of imaging and back-scattered photons (temporal impulse response) are plotted. It is clear from this figure that backscattered photons arrive much earlier than image photons. Fig. 6.11 and Fig. 6.12 show the simulation results for an optical thickness 5 (10 round trip). In this case, most of the received photons are the backscattered ones. However, since there is a difference in the arrival time of image and back-scattered photons, there is still hope for seeing through cloud if we have enough power.

As an example, a femtosecond pulsed laser that operates with 50fs pulses at 800nm wavelength and a repetition rate of 3 gigapulses per second with an average power of 1W, gives a peak power of 6.67kW transmitted. This translates to a transmitted energy of 0.3nJ per pulse. For an optical thickness of 1, Beer-Lambert's law dictates that the attenuation experienced by the line-of-sight (LOS) laser beam on its way to the target and back to the image plane will be $e^{-2\tau}$, which is 0.1353, and the received LOS pulse will have an energy of 0.0406nJ, which is too small to be detected by current state-of-the-art optical receivers in the presence of background noise. However, decreasing the repetition rate to 3 megapulses per second, with the other parameters kept fixed, would increase the received pulse energy to 0.0406μJ, a 1000-fold improvement over the previous repetition rate. For an optical thickness of 5, the aggregate roundtrip thickness experienced by the beam is 10, and the attenuation is 4.54×10^{-5}. Then the received energy per pulse for 3 gigapulses per second becomes 1.362×10^{-5} nJ, and for 3 megapulses per second it becomes 1.362×10^{-5} μJ. An examination of these values indicates that we may not expect a laser imaging system to operate above the noise floor for clouds having high optical thicknesses. For a MIMO system to yield some form of diversity in clouds without scintillation, we would need higher optical thicknesses so that the light energy is dispersed and then captured by receivers corresponding to transmitters other than the intended. But going to such high thicknesses seems impractical in terms of detectability, and we must bank on scintillation only for obtaining the performance benefits inherent in MIMO.
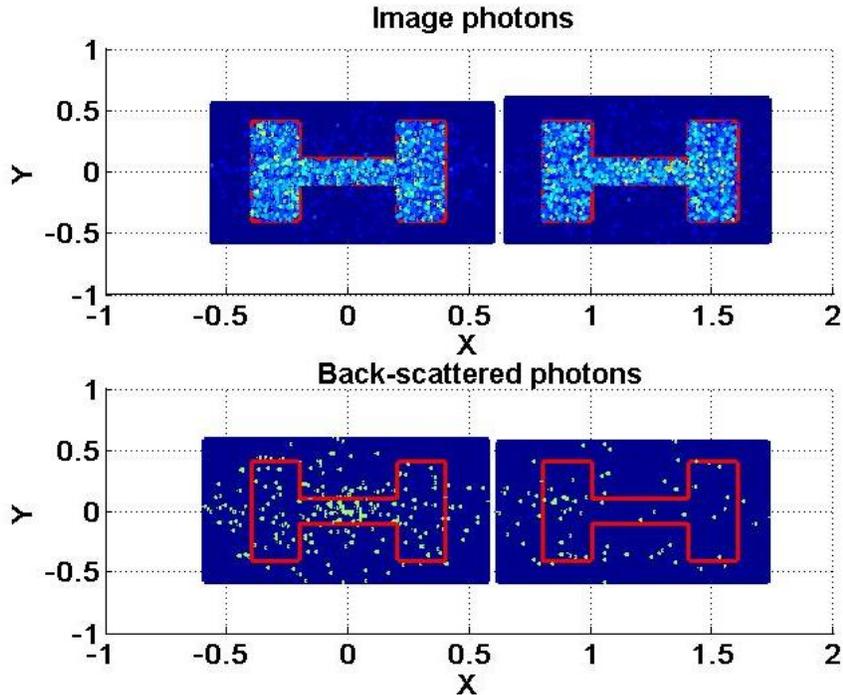
Fig. 6.9: Imaging and back-scattered photons for a 2x2 MIMO imaging system working in a medium of optical thickness 2.
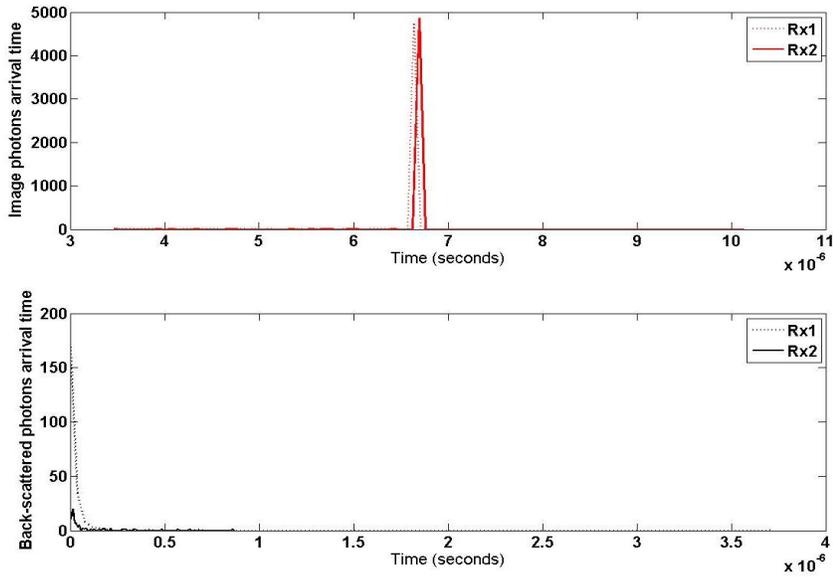
Fig. 6.10: Arrival times of imaging and back-scattered photons for a 2x2 imaging system operating in a medium of optical thickness 2.
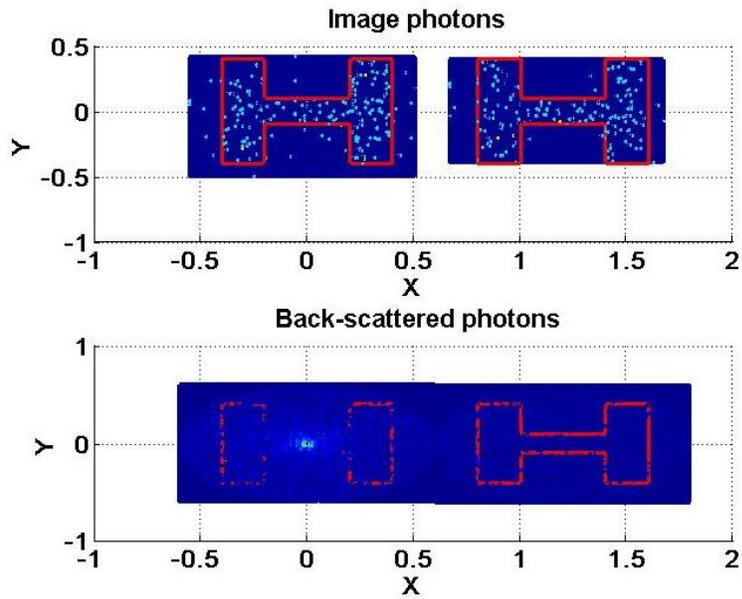
Fig. 6.11: Imaging and back-scattered photons for a 2x2 MIMO imaging system working in a medium of optical thickness 5.
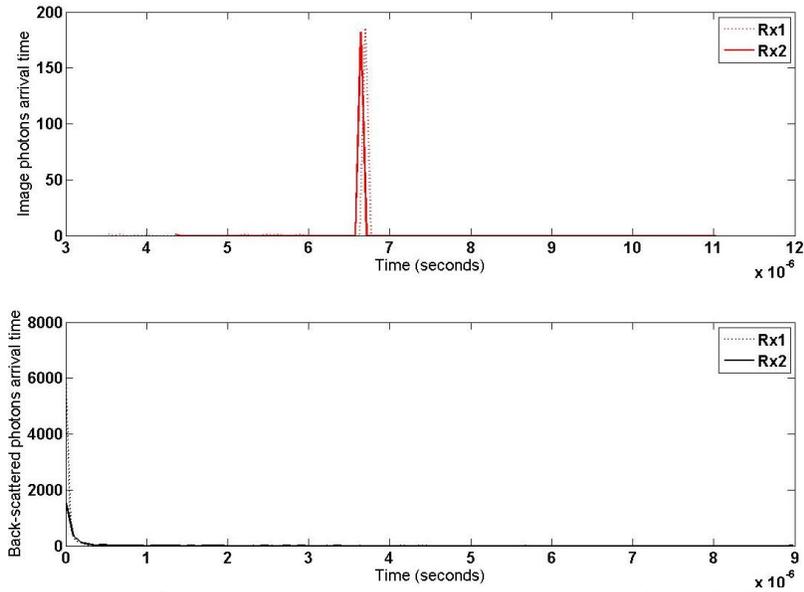
Fig. 6.12: Arrival times of imaging and back-scattered photons for a 2x2 imaging system operating in a medium of optical thickness 5.

6.3   Transmitter Design

For imaging purposes, we can resort to two widely used methods of obtaining reflections from a target: (1) "flood" or completely illuminating the target with a broad laser beam, and (2) scanning the target by a high-power spatially-concentrated beam, which is also known as "raster scanning". Both of these methods have their own shortcomings.

A Gaussian laser beam does not uniformly illuminate the target. As a result, the Signal-to-Noise Ratio (SNR) per pixel would not be the same for all portions of the image. Moreover, the spreading of the laser beam will significantly reduce the intensity even in clear weather condition. This attenuation is more severe in presence of cloud and turbulence. Furthermore, distorted point spread function introduces random interferences between neighboring pixels, and affects resolution and contrast performances.

Raster scanning, on the other hand, suffers less from attenuation because the laser beam is confined in space. However, since the entire target needed to be scanned to form the image, the area search rate is low for this approach. This method is widely used in LIDAR applications in clear weather conditions.

One intermediate approach is to split the Gaussian laser beam into an array of beams of smaller waists or beamlets. Photolithographic beam-splitters can convert a transmitted Gaussian laser beam into an array of $N \times N$ quasi-uniform intensity spots in the far field. There are two key parameters in designing the beam splitters: uniform distribution of intensity or power between the beamlets, and efficiency of converting Gaussian beam into multiple beams. Fig. 6.13 shows a *10 × 10* array of uniform intensity spots, produced by a photolithographic beam-splitter [7]. These beamlets are separated in space by a suitable distance to be substantially non-interfering and are pointed to different portions of the target for uniform illumination. This method is preferred over floodlit illumination due to better geometric coupling, reduced background radiation, and increased angular resolution [50]. In

other words, we would like to have a transmitter that produces multiple beams with prescribed intensities covering the target area.
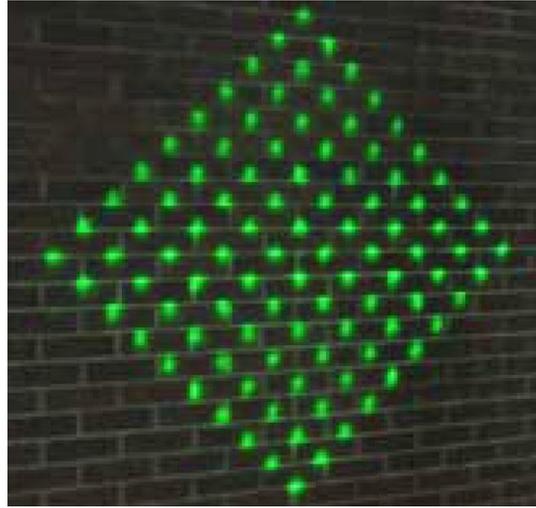
Fig. 6.13: Array of 10 x 10 uniform intensity spots produced using a photolithographic beam-splitter.
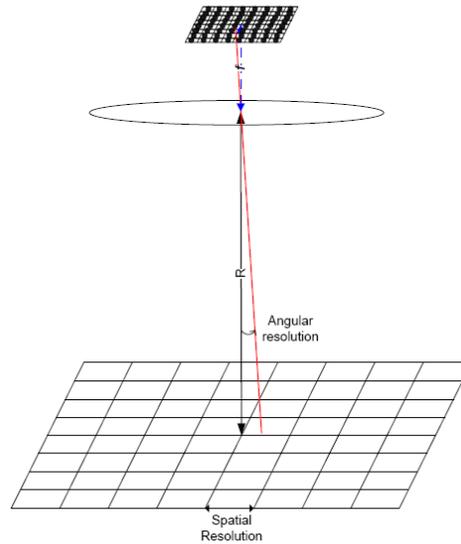
Fig. 6.14: Simplified view of the imaging system.

Imaging photolithographic optical elements have been used at the transmitter to generate multiple beams from a single laser diode [7][52][53][54]. Holograms generated by means of a computer can produce wave-fronts with any prescribed amplitude and phase distribution. Computer generated holograms (CGH) [53] have many useful properties. An ideal wave-front can be computed on the basis of diffraction theory and encoded into a tangible hologram. A multilevel phase CGH can have diffraction efficiency close to 100%. Holographic Optical Elements have insignificant physical weight and are low-cost when mass produced; hence, being specially suited for on-board of aircrafts and spacecrafts. In principle, the more the diffusing spots are, the more optical power can be transmitted in compliance with the eye

safety regulations. In addition, a more uniform distribution and a higher predictability of the channel parameters can be achieved. The beams emerging from a hologram are almost collimated. Increasing the number of diffusing spots will allow increasing the intensity of each spot and of the total transmitted optical power. This is a very important aspect of photolithographic beam-splitters as increasing the total transmitted power enables the system to operate in larger optical depths. To maximize the signal optical power received by each of the branches, each diffusing spot should contain as much energy as possible. Let us assume that for an array of $N \times N$ beamlets, the central spot contains 1% of total transmitted optical power, P_t, i.e., $P_c = P_t / 100$, and the CGH is designed so that the central spot is one of the spots in the array. Then $P_t = (N \times N - 1) P_d + P_c = 100 P_c$, where P_d is the power contained in each of diffusing spots, and $N \times N$ is the total number of spots. Each of the collimated beams emerging from the hologram should contain no more energy than $kAEL$, where AEL is the Accessible Emission Limit for a point source, established by the International Standard IEC 825 [55], and $k = (D_{beam} / D_{aperture})^2$, D_{beam} being the beam diameter and $D_{aperture}$ being the aperture diameter applicable to measuring laser irradiance and radiant exposure [55]. Thus, maximum total power is transmitted when the central spot energy equals the accessible emission limit: $P_c = kAEL$. Then, the power contained in each of the other diffusing spots is $P_d = \dfrac{99}{N \times N - 1} kAEL \leq kAEL$. The maximum for P_d, i.e., $P_d = kAEL$, is achieved for a number of spots $N \times N = 100$. In our case study we consider a square image, i.e., a spot array of 10×10 spots would be the optimum. Fig. 6.15 shows a sample photolithographic 4×4 beam splitter that we have produced. One can infer the small size and the light weight of this beam splitter from Fig. 6.15. Using a 532 nm low power laser light (<5 mW) through the sample, we captured the spots on a camera as demonstrated in Fig. 6.16.
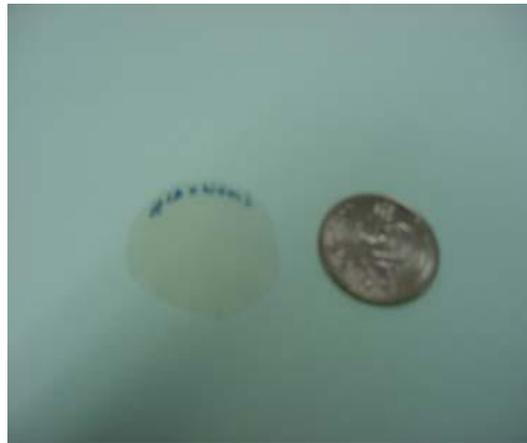
Fig. 6.15: Sample plastic 4x4 photolithographic beam splitter on the left of a quarter.
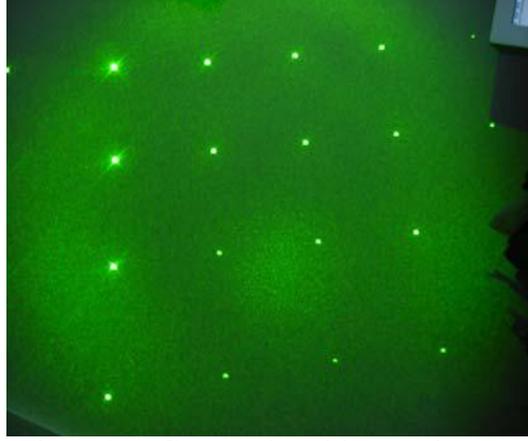
Fig. 6.16: Captured multi-spots image.

The initial spreading angle of beamlets is set such that they do not interfere with one another on the target plane. For a Gaussian beam of initial spreading angle θ_0, the initial beam waist is $\omega_0 = \lambda / \pi\theta_0$, where λ is the beam wavelength. Beam waist at a distance R from the source, $\omega(R)$, is related to ω_0 as:

$$\omega(R) = \omega_0 \sqrt{1 + \left(\frac{\lambda R}{\pi\omega_0^2}\right)} \qquad (6.3)$$

For an imaging system operating at a wavelength of $\lambda = 1.55$ μm, and a target in $R = 1$ km, spatial resolution is limited by beam waist at the target, $\omega(R)$. For example if a spatial resolution of *20 cm* is required, $\omega(R)$ should be less than *20 cm*, which means an initial spreading angle of less than 200 μrad.

6.4   Receiver Design

An *N×N* array of photo-detectors is used to receive the reflections from the target. Fig. 6.17 shows a schematic of the proposed system. Typically, the optical front-end of the receiver consists of an optical concentrator to increase the received optical signal power, and an optical band pass filter to reject the ambient light. We have investigated the characteristics and performance of optical interference filters. Typically, interference filters with a narrowband spectral response have a low transmittance and reduce the signal by 3 dBo. These filters work well with collimated radiation which is not an option in our application. The narrowband interference filters are sensitive to incident angle of radiation and the peak wavelength shifts towards shorter wavelengths with increasing the incident angle. The filter can be designed with a shift towards longer wavelengths in order to accommodate the spectral shift by an incident angle.
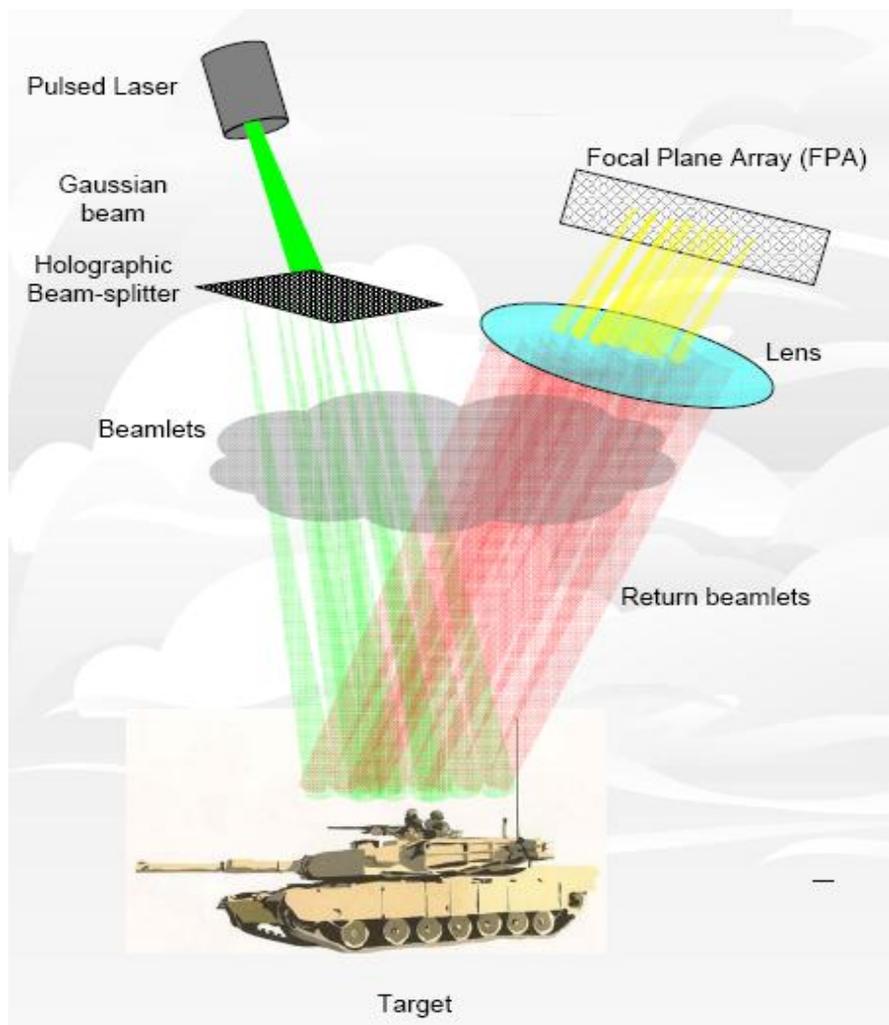
Fig. 6.17: Simplified schematic of the proposed imaging system.

Given a diffuse radiation pattern, however, performance further deteriorates because of non-collimated radiation, i.e., the incident angle varies from normal incidence to the maximum angle determined by the receiver lens FOV. When the spectral response of the filter is averaged over all incident angles, the full-width-half-maximum (FWHM) is broadened and the signal is further attenuated. The maximum transmittance is further reduced for non-polarized light; the case in imaging through cloud/fog. Since the laser diode and the filter wavelengths are specified by the manufacturer within given tolerances, and there is also a wavelength drift due to temperature changes, a narrow-band interference filter does not seem to be practical. Using the unique properties of holographic optical elements (HOE), a novel design is proposed for the receiver optical subsystem used in the proposed imaging system. With a holographic curved mirror as an optical front-end, the receiver would achieve several dBs of improvement in the electrical signal-to-noise ratio compared to a bare photo-detector. The main advantages of HOEs over conventional systems that consist of a lens concentrator and an optical filter are multi-functionality, independence of their physical configuration, insignificant weight, and low cost.

The filtered and concentrated beamlets are then concentrated onto an $N \times N$ array of high-speed photo-detectors. Examples of high-speed photo-detectors suitable for precise multi-

channel ranging and imaging include segmented anode photomultiplier tubes (PMT), and avalanche photodiodes (APD). Fig. 6.18 shows a sample 32×32 APD array [8][9][10].
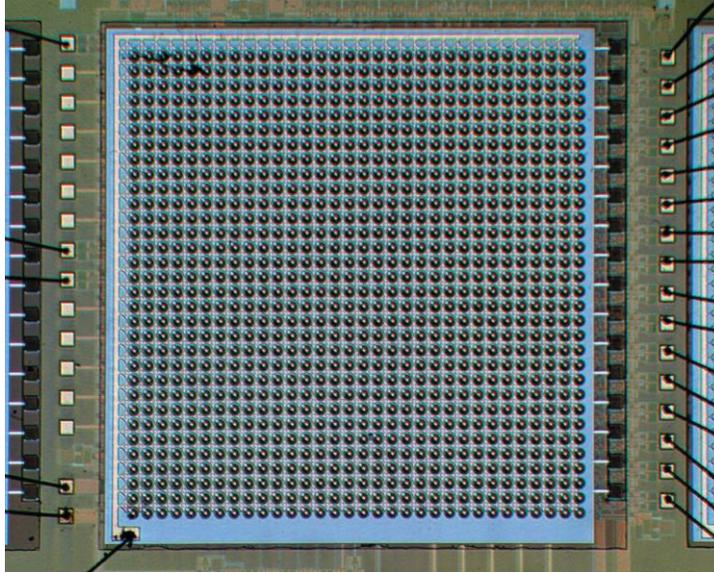
Fig. 6.18: Photomicrograph of a bridge-bonded APD/CMOS device showing the 32×32 array.

Assuming that each photo-detector has a diameter of 100 μ m (30 μm active area diameter) [50] and lens has a focal length f, the angular resolution of the system would be $100/f$ μ rad. This translates into a spatial resolution of $R \times 10^{-4}/f$, where R is the target range. Moreover, the receiver FOV turns out to be $N/f \times 100$ μrad. As an example, suppose that we have a *10×10* array of beamlets imaging an object which is *1 km* away. Furthermore, suppose we have a receiving aperture of *D=50 cm* diameter with *f=D=50 cm*. The angular resolution of this system would be *2 m rad*, which translates to spatial resolution of *20 cm*. Hence, a rectangle of *4 m²* can be imaged by this system. Fig. 6.14 demonstrates a simplified geometry of the imaging system. In practical systems, the same aperture is used for both reception and transmission. While beamlets are transmitted from central part of the aperture, return beams are received by peripheral area of the lens.

6.5   Power-Energy Analysis

The power requirements of an optical imaging system depend on the path loss, receiver aperture area, and the photo-detector sensitivity. As an example we assume a 250 ps pulsed laser with 3 μJ energy per pulse and a repetition rate of 1 kHz, operating in 532 nm wavelength [8][9][10]. While the peak power of this laser is about $3\mu J / 250 ps = 12$ k watts or 40.8 dB, the average power is about $3\mu J / 1ms = 3$ m watts or 5.8 dBm.

The energy of each photon is given by $E = h\nu = 3.73 \times 10^{-19}$. This means that each pulse approximately contains 8×10^{12} photons. Assuming a channel of optical thickness 5, the round trip attenuation of each pulse is $e^{-10} = 4.54 \times 10^{-5} = -43.42$ dB. Hence, the received power is about 27.4 dBm. Furthermore, the number of received photons is 3.6×10^{8}.

If one considers all the degrading factors in the link budget, the received signal measured in the number of detections or photoelectrons, N_{det} is given by [50]:

$$N_{\det} = \frac{E_p}{h\nu} \tau_{to} \tau_{ro} \tau_{atm}^2 \frac{1}{N_{pix}} \frac{A_{CA}}{\pi R^2} \eta_{\det} f_{tr} f_{ll} \qquad (6.4)$$

where:

E_p : Energy Per Pulse,

$h\nu$: Energy Per Photon,

τ_{to} : Transmission of transmit optics,

τ_{ro} : Transmission of receive optics,

$\tau_{atm} = e^{-\tau}$: Transmission of atmosphere,

N_{pix} : Number of pixels,

A_{CA} : Collecting Area,

R : Range to the target,

$\eta_{\det}$: Quantum efficiency of detector,

f_{tr} : Overlap of Tx beam and Rx FOV,

f_{ll} : Efficiency of optical coupling.

The state of the art technology provides us with photon-counting receivers with 0.5 photons per bit sensitivity. There are also Low Light Level Charge-Coupled Devices (L3CCDs) with ability of accurate estimation of signals as faint as 1 photon per pixel per frame [56]. Imaging through foliage and trees has been reported using these arrays of APD photo detectors [8], which basically measure the arrival times of the photons rather than the photon count. Given the above analysis and literature review, it can be stated that the state of the art technology may enable imaging through clouds.

6.6   MIMO Imaging in Turbulent Atmosphere

Under cloudless, turbulence-free atmospheric conditions, an $N \times N$ array of Gaussian beamlets is launched towards the target. Since there is no scattering and scintillation in such a scenario, beamlets arrive at the target with no distortion and interference with one another. At target plane, beamlets are amplitude modulated with reflection coefficients of the corresponding areas and reflected back to the receiving aperture. Received beamlets are then concentrated on photo-detectors. Most prior works have investigated the MIMO imaging performance for relatively short link lengths and under such an ideal atmospheric condition [50][8].

In Section 6.2, it is clarified that the effects of clouds and other scattering media can be summarized in an attenuation factor, since after time-gating, only the coherent component of the scattering impulse response contributes to the total channel impulse response. In this section, attributes of turbulence on the imaging system performance is investigated. Under turbulent atmospheric condition, transmitted beamlets distort and may interfere with one another on their way to target. The same problem exists in return path from target to receiver. Using a computer-generated hologram (CGH) [53], it is possible to pre-distort the transmitted beamlets to overcome atmospheric turbulence in "down-link" path. To compensate for

atmospheric turbulence in "uplink", one has to first model scintillation. In chapter 5, we described different methods of generating phase-screens which are used to approximate phase perturbations. These phase perturbations distort and broaden the Point Spread Function (PSF), which plays the role of impulse response in imaging systems. In other words, PSF is convolved with focal-plane signal and the result is direct-detected by photo-detectors. Due to PSF broadening and distortion, information from neighboring pixels may mix and Inter-Pixel-Interference (IPI) is observed. Leakage of information into neighboring pixels gives rise to IPI. In other words, PSF becomes so broad that mixes the intensity of neighboring elements. The mathematical formulation for a linear degradation caused by blurring and additive noise is given by:

$$g(\mathbf{x}) = (f * h)(\mathbf{x}) + n(\mathbf{x}) = \sum_{s \in S_h} h(\mathbf{x} - \mathbf{s}) f(\mathbf{x}) + n(\mathbf{x}) \tag{6.5}$$

where $f(\mathbf{x})$, $g(\mathbf{x})$, $h(\mathbf{x})$, and $n(\mathbf{x})$ represent the original image, the observed image, the blur or PSF, and the observation noise, respectively, $S_f \subset \mathfrak{R}^2$ is the image support, and $S_h \subset \mathfrak{R}^2$ is the PSF support. Support refers to the smallest rectangle within which the true object is contained. In the case of coherent, monochromatic illumination, the following simple relationship exists between the object field, $u_o(\mathbf{x})$, and the image field, $u_i(\mathbf{x})$;

$$u_i(\mathbf{x}) = u_o(\mathbf{x}) * h(\mathbf{x}) \tag{6.6}$$

where $h(\mathbf{x})$ is the impulse response of the imaging system and is given by:

$$h(\mathbf{x}) = \int W(f \lambda d_i) \exp(-j2\pi\mathbf{f}.\mathbf{x}) \times \exp(i\varphi) d\mathbf{f} \tag{6.7}$$

where $W(f \lambda d_i)$ is the pupil function of the imaging system, λ is the optical wavelength, and d_i is the distance between the exit pupil and the image plane. Furthermore, φ represents the phase front distortion caused by atmospheric turbulence. From (6.6) one can see that (6.5) clearly describes the image recovery problem in the case of coherent imaging. For incoherent imaging, however, the irradiance observed in the image plane is given by the squared modulus of (6.6) as:

$$i(\mathbf{x}) = |u_i(\mathbf{x})|^2 = |u_o(\mathbf{x}) * h(\mathbf{x})|^2 \tag{6.8}$$

Expanding (6.8) gives:

$$i(\mathbf{x}) = \int \int u_0(\mathbf{x}') u_0^*(\mathbf{x}'') h(\mathbf{x} - \mathbf{x}') h^*(\mathbf{x} - \mathbf{x}'') d\mathbf{x}' d\mathbf{x}'' \tag{6.9}$$

Due to atmospheric effects, we can assume that the object field $u_o(\mathbf{x})$ is inherently a random quantity, which in turn implies the irradiance is also random. Hence, the average irradiance observed in the imaging plane becomes:

$$\langle i(\mathbf{x}) \rangle = \int \int \langle u_0(\mathbf{x}') u_0^*(\mathbf{x}'') \rangle h(\mathbf{x} - \mathbf{x}') h^*(\mathbf{x} - \mathbf{x}'') d\mathbf{x}' d\mathbf{x}''$$
$$= \int \langle o(\mathbf{x}') \rangle |h(\mathbf{x} - \mathbf{x}')|^2 d\mathbf{x}' = \langle o(\mathbf{x}) \rangle * |h(\mathbf{x})|^2 \tag{6.10}$$

where $\langle o(\mathbf{x}) \rangle$ is the average intensity of the object. The brackets can be dropped for notational simplicity. The impulse response characterizing the input/output relationship between the two irradiances is called the point spread function and is given by:

$$s(\mathbf{x}) = \mid h(\mathbf{x}) \mid^2 \qquad (6.11)$$

It is interesting to note that, in the case of incoherent illumination, the imaging system responds linearly to the object irradiance distribution, where as for coherent illumination, the system responds linearly to the object field distribution [6].

The Full Width Half Maximum (FWHM) of PSF is usually used as a measure of resolution in most imaging systems. When PSF is an ideal airy pattern, diffraction-limited imaging is possible. As turbulence level increases, PSF becomes broader and more distorted and resolution decreases, regardless of the number of beamlets and their compact placement.

Fig. 6.19 shows the focal plane image for an *8×8* MIMO imaging system under turbulence-free conditions, where the object to be imaged is letter "H". Fig. 6.20 shows the results under moderate turbulence. These figures have been obtained assuming a channel length of *L=5 km* (one way), $C_n^2(0) = 1.7 \times 10^{-12}$, $r_0 = 1.4$ *cm*, and a receiving aperture of *50 cm* diameter.
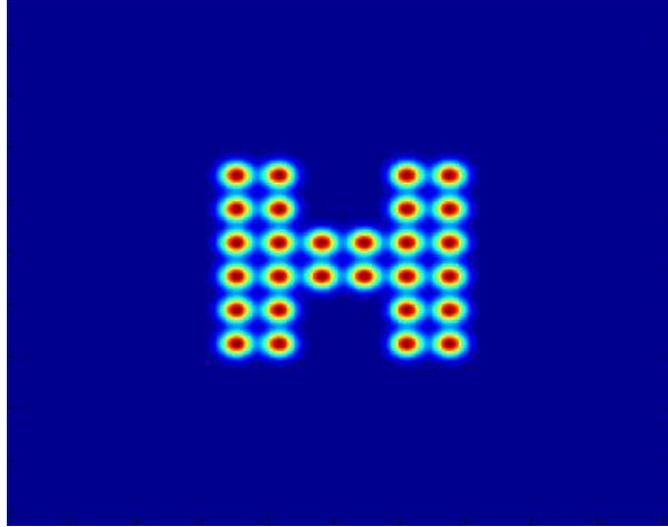
Fig. 6.19: Focal plane wave for turbulence-free condition, where the object is letter "H".
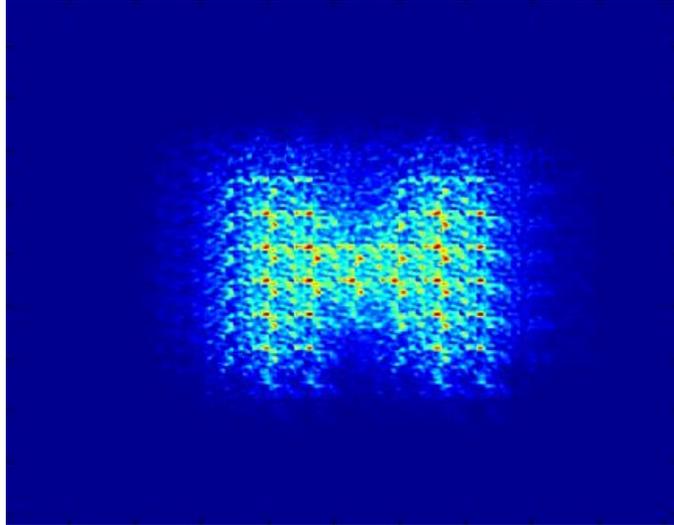
Fig. 6.20: Focal plane wave for moderate turbulence, where the object is letter "H".

Fig. 6.21 and Fig. 6.22 show the photo-detected images corresponding to turbulence-free and moderate turbulence, respectively.
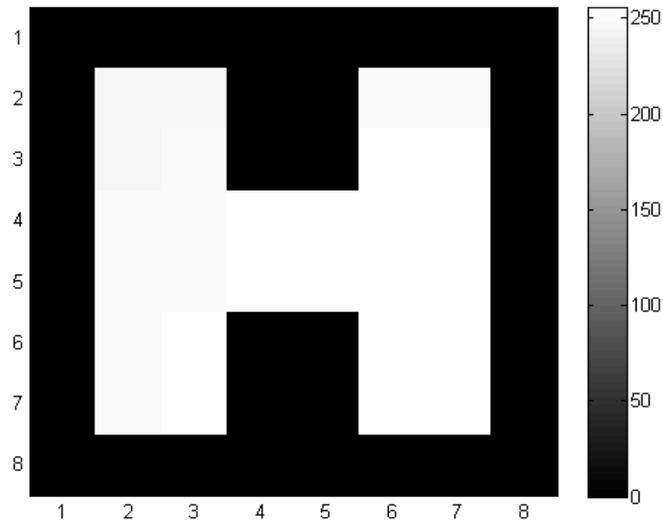
Fig. 6.21: Photo-detected image under turbulence-free condition.
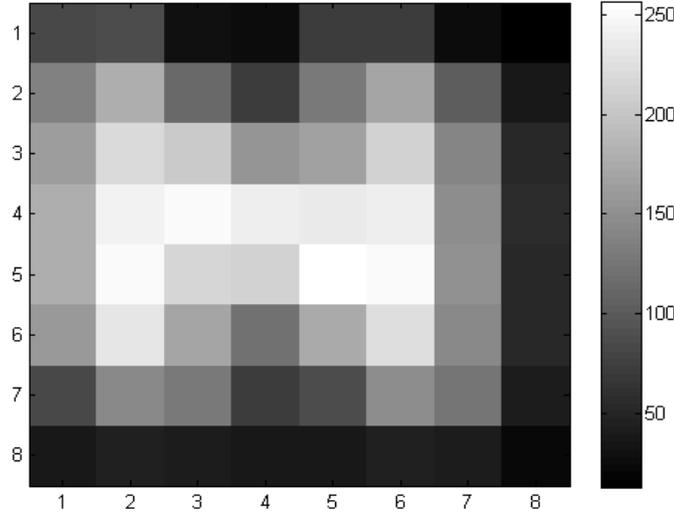
Fig. 6.22: Photo-detected image under moderate turbulence.

6.7   *A More Accurate Channel Model*

In this section details of a more realistic channel model is described, which exploits multiple phase screens in the path of laser beamlets [43]. Fig. 6.23 demonstrates the channel model configuration. In Fig. 6.23, three equidistant (ΔL) phase screens are used throughout a channel of length L to simulate the effects of turbulence. Note that only the forward path is shown in this figure. A laser beamlet is launched by the transceiver and is Fourier-propagated to the first phase screen. Assuming the initial beamlets to have a spatial form of $t_0(x, y)$, the wave impinging on the first phase screen is obtained using Fourier optics as:

$$
\begin{aligned}
t_1(x, y) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} T_0(f_x, f_y) H(f_x, f_y, \Delta L) e^{j2\pi(f_x x + f_x y)} df_x df_y \\
&= F^{-1}\left(F\left(t_0(x, y)\right) \times H(f_x, f_y, \Delta L)\right)
\end{aligned}
\tag{6.12}
$$

where $t_1(x, y)$ is the spatial domain expression for the beamlet impinging on the first phase screen, $T_0(f_x, f_y)$ is the Fourier transform of $t_0(x, y)$, and $H(f_x, f_y, \Delta L)$ is the transfer function of free space propagation and is given by:

$$
H(f_x, f_y, \Delta L) = \exp\left(j2\pi \sqrt{\frac{1}{\lambda^2} - f_x^2 - f_y^2} \right)
\tag{6.13}
$$

Note that (6.13) is the exact expression for free space transfer function assuming scalar diffraction theory. It is also possible to apply paraxial (small angle) or far field approximations and simplify this equation based on Fresnel or Fraunhofer diffraction theories [57].

When beamlet impinges on the first phase screen, wave-front is perturbed and then beamlet is Fourier-propagated to the second phase screen. The wave impinging on the second phase screen is obtained using Fourier optics as:

90

$$t_2(x, y) = F^{-1}\Big(F\big(t_1(x, y) \times e^{j\phi_1(x,y)}\big) \times H(f_x, f_y, \Delta L)\Big) \tag{6.14}$$

where $\varphi_1(x, y)$ is the first phase screen, and $t_1(x, y)$ is the spatial domain expression for the beamlet impinging on the first phase screen. Using a similar approach, beamlet propagation from the transceiver to the target and back is simulated. Note that beamlets pass through the same set of phase screens on both downlink and uplink.

Fig. 6.24 (a), (b), and (c) show the focal plane image of an "H"-shaped target for weak, moderate, and strong turbulence conditions, where the one way channel length is 3 km and five phase screens are placed through the channel. Furthermore, wind speed is set to 21 m/s and the structure parameter of refractive index variations, $C_n^2(0)$, is set to 1.7×10^{-13}, 6×10^{-13}, and 1.7×10^{-12} for relatively weak, moderate and strong turbulence conditions, respectively. This channel model can produce more accurate results, but at the same time it is more complicated to analyze. In other words, in contrast to the single phase screen channel which distorts all the beamlets almost the same, this model induces different aberrations on beamlets depending on their direction and appears to be linear but shift variant. Moreover, due to constructive and destructive interference of beamlets with themselves, one can observe scintillation effects.
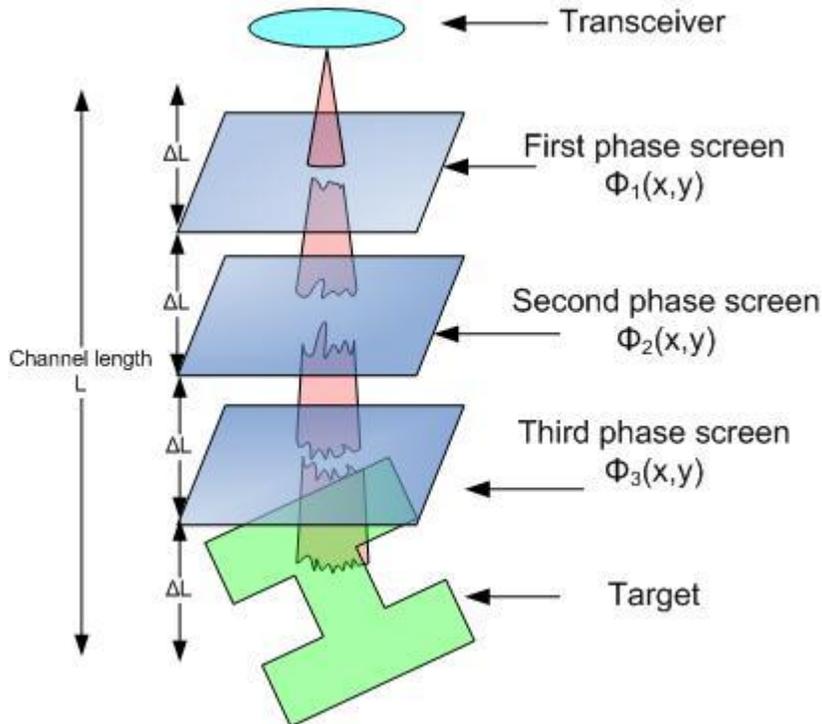
Fig. 6.23: Optical channel model configuration.
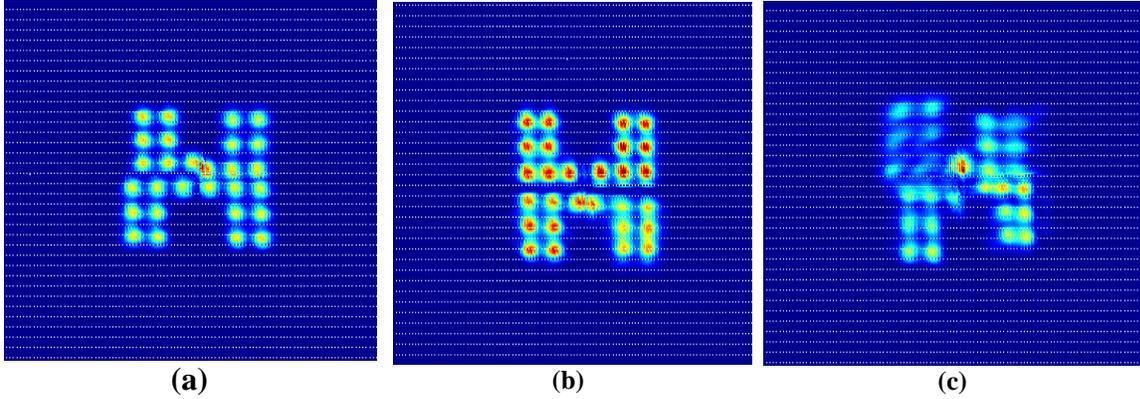
Fig. 6.24: Focal plane image for (a) weak, (b) moderate, and (c) strong turbulence conditions.

6.8   *Conclusions*

In this chapter, the overall design of a spatially multiplexed MIMO imaging system is described. Moreover, attributes of turbidity and turbulence in the propagation media and their impact on imaging system are briefly investigated. While scattering media give rise to steady background noise via backscattering and severely attenuates the information-bearing coherent component of the received wave, according to beer-Lambert law, turbulence contributes to amplitude and phase fluctuations, and hence, causes PSF broadening and signal fading. In order to restore the received image to its ideal form, the imaging system should be well-equipped with Adaptive Optics (AO), proper spatial and temporal filtering and powerful post-processing modules such as Blind Deconvolution (BD) blocks. In the next chapter, we will discus performance improvements obtained via embedding these blocks and modules.

Chapter 7
Compensation and Restoration Techniques & Performance Improvements

7.1   *Introduction*

In chapter 6, a spatially multiplexed optical MIMO imaging system is introduced and transceiver design is illustrated. This system is proved to work properly under ideal atmospheric conditions, where effects of scattering and turbulence are minimal [50][8][9][10]. However, as shown in chapter 6, in imaging through atmosphere one should consider dispersive effects of scattering particles, turbulence-induced beam spreading, intensity fluctuations, and wave-front distortion. While scattering media give rise to steady background noise via backscattering and severely attenuate the information-bearing coherent component of the received wave, according to Beer-Lambert law, turbulence contributes to amplitude and phase fluctuations, and hence, causes PSF broadening and signal fading.

In order to overcome the deleterious effects of scattering and remove the clutter sub-space (steady background noise due to backscattering) from image sub-space, one can resort to some kind of spatio-temporal filtering, and take advantage of the fact that spatial and temporal distribution of image and backscattered photons are approximately non-overlapping. In other words, one can expect the image photons to arrive at the receiving aperture in a predictable time interval, and within a limited FOV. In this chapter, we use time-gating to filter-out the backscattered photons. It is shown that SNR or Contrast can be improved drastically, depending on the length of the time-gate and the optical thickness of scattering medium.

Restoring the effects of turbulent atmosphere is complicated by the fact that imaging system does not have any information about either the true image or the distorting PSF. Solutions to this problem fall into two major categories; Adaptive Optics corrections and deconvolution Algorithms. While Adaptive Optics systems estimate the wave-front phase aberrations and compensate for PSF broadening in real-time, blind deconvolution algorithms serve as a post processing module and try to recover both PSF and true image from the received distorted image.

In the following sections, we elaborate on image restoration techniques and the resulting performance improvements in terms of Signal-to-Noise Ratio Improvement (SNRI), Signal-to-Noise-and-Interference Ratio (SNIR) or Contrast, Resolution, and Mean-Square Error (MSE) percentage.

7.2   *Clutter Space Removal via Time-Gating*

In this section, we revisit chapter 6, where it is shown that time gating can improve SNIR or contrast by filtering out backscattered photons from image information bearing photons. Knowing that received energy is proportional to the number of photons, Signal to Noise and Interference Ratio (SNIR) can be defined as:

$$SNIR = 10\log_{10}\left(\frac{N_I h\lambda}{N_b h\lambda + N_n h\bar{\lambda}} \right)$$

(7.1)

where N_I is the number of image photons, N_b is the number of backscattered photons, N_n is the number of noise photons, h is Planck's constant, λ is the wavelength of imaging system carrier, and $\bar{\lambda}$ is the average wavelength of noise photons. Assuming that backscattered photons are dominant in the steady background noise, the second term in denominator of *log* in (7.1) can be omitted, and hence the term in the *log* can be simplified to the ratio of number of image and backscattered photons. We run the MCRT imaging code developed at Penn State CICTR for various optical thickness values and record the number of received image and backscattered photons. The ratio of these two numbers determines the SNIR or contrast before time gating. Then by applying time gating, the number of photons arrived at the receiver in a specified time-frame about the round-trip time is measured. Table 7.1 shows the results of this analysis for different optical thickness values (one way) and different time-gate lengths.

Table 7.1: Contrast Improvement using time-gates of different lengths.

Optical Thickness /SNIR	1	2	3	4	5	6
before Time-Gating	26.3 dB	14.5 dB	4.0 dB	-5.5 dB	-15 dB	-25.2 dB
10 µs Time-Gating	49.9 dB	41.3 dB	33.2 dB	26.1 dB	15.4 dB	6.3 dB
1 µs Time-Gating	49.9 dB	49.8 dB	42.5 dB	43.3 dB	34.6 dB	16.2 dB
0.1 µs Time-Gating	49.9 dB	49.8 dB	48.5 dB	43.3 dB	34.6 dB	24.6 dB

From Table 7.1, one can see that SNIR decreases with optical thickness in all cases. This is due to the fact that as optical thickness increases, fewer number of image photons can make it to the receiver. Moreover, probability of backscattering increases as more scattering events take place in the medium. By placing a time-gate of 10 µs length with its center on the round-trip time, SNIR improves, significantly. This improvement is even more considerable for larger optical lengths. By shortening the time-gate to 1 µs, we observe more improvement, especially in larger optical thickness values. If we further decrease this length to 0.1 µs, the SNIR improves only for a channel of optical thickness 6 (12 round-trip). These final SNIR improvements with a gate of 0.1 µs length appear to be very promising, as we observe a total of 50 dB improvement for a channel of optical thickness 6. In order to continue this analysis for larger optical thickness values, MCRT needs a very long processing time. This is due to the fact that signal attenuation increases exponentially with optical thickness, and hence a larger number of photons need to be tracked, in order for the receiver to get some information bearing photons.

7.3   *Combining Multiple Images*
In this section, it is shown how combining multiple shots of the same object may help us to overcome PSF distortions. Note that, this method is useful when receiver has no knowledge of PSF and IPI is not severe. In this case, by averaging multiple images, one can compensate for beam wandering over the surface of photo-detector. However, if information from the neighboring pixels mix completely and spatial "eye closure" happens, this method is no longer helpful.

Under weak turbulence conditions, PSF is distorted and broadened; however, this broadening is not severe enough to cause IPI. As a result, one can obtain an image of good quality by averaging several distorted images. Fig. 7.1 shows a photo-detected image under weak turbulence condition and Fig. 7.2 shows an image obtained by averaging 20 distorted

images under the same condition. One can observe that averaging can virtually reverse the turbulence effects.

In moderate turbulence conditions, similar to that of Fig. 7.3, PSF is completely distorted but the mixing between neighboring pixels is still tolerable. In this case, if multiple images are obtained in some uncorrelated time intervals and added up with equal weights, we expect the resulting image to show less variation. Fig. 7.4 shows the image resulted from averaging 20 distorted images in moderate turbulence conditions similar to Fig. 7.4. It is clear that image has improved, significantly. Furthermore average contrast is increased compared to Fig. 7.3. However, comparing Fig. 7.2 and Fig. 7.4, we realize that by increasing the turbulence level, the averaging method becomes less effective.
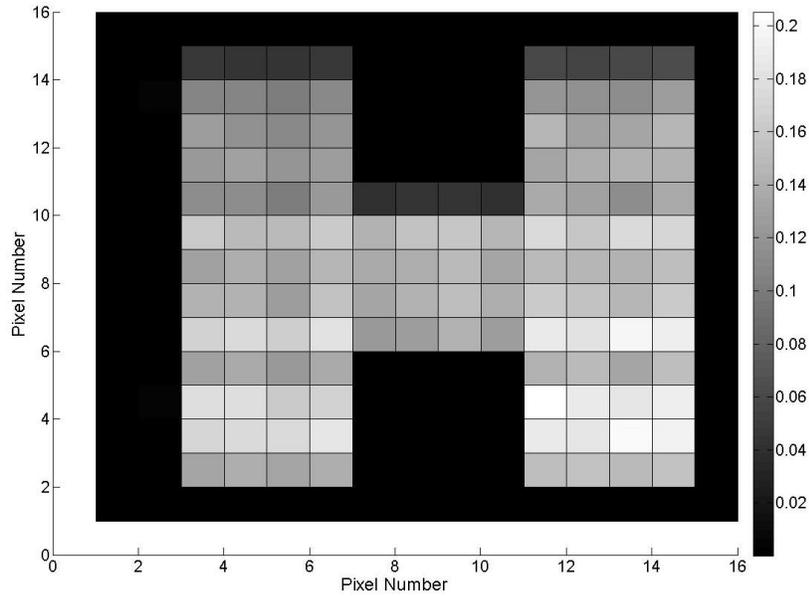
Fig. 7.1: Photo-detected image under weak turbulence.
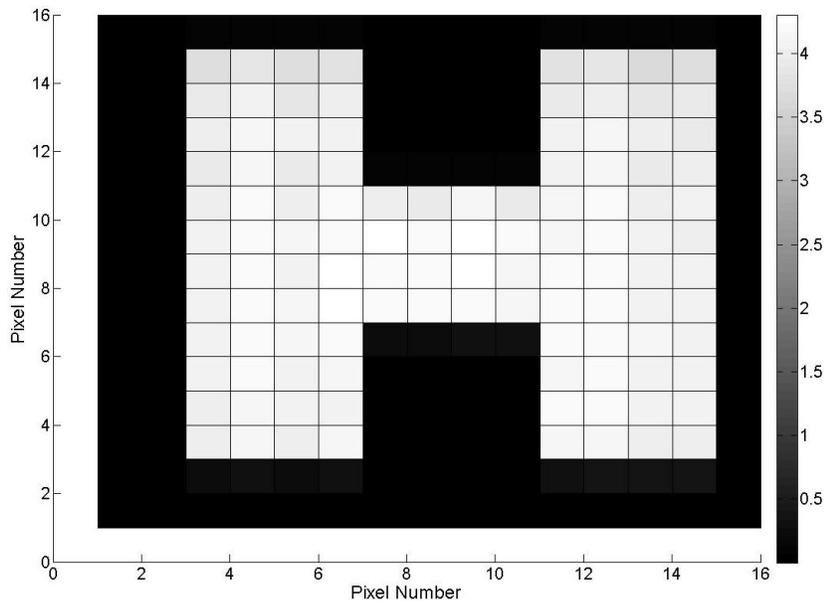
Fig. 7.2: Image resulted from averaging 20 distorted images under weak turbulence
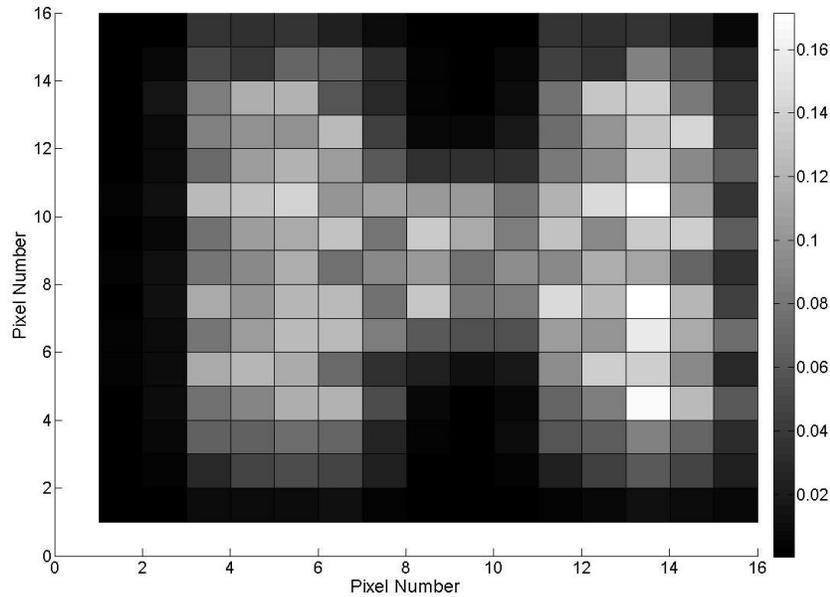
Fig. 7.3: Photo-detected image under moderate turbulence.
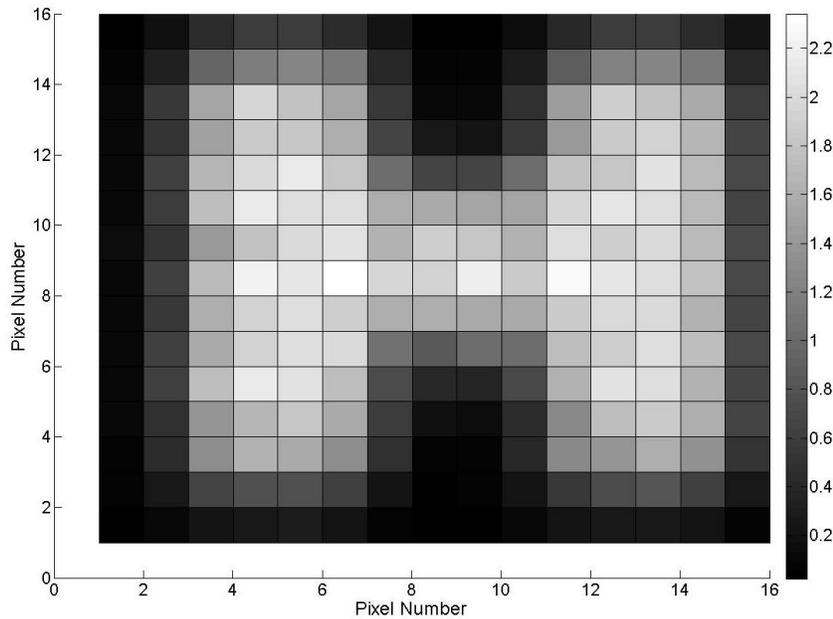
Fig. 7.4: Image resulted from averaging 20 distorted images under moderate turbulence

7.4 *Deconvolution Techniques*

In order to overcome the blurring distortion caused by the non-ideal PSF induced by turbulence, the distorted PSF must be deconvolved from the image. However, this is a very complicated problem, since we have to estimate both target and PSF from a distorted image. If PSF is available, image may be recovered using a variety of well known techniques, such as inverse filtering, Wiener filtering, least square (LS) filtering, recursive Kalman filtering, and constrained iterative deconvolution methods [3]. Otherwise multiple images from a same

object must be combined to remove random distortions. Next we will describe these retrieving methods.

7.4.1   Inverse Filtering

If perfect Channel Side Information (CSI) is available, original image can be recovered easily. Using MIMO formulation, distorted image can be written as:

$$\mathbf{y} = \mathbf{Rx} \qquad\qquad (7.2)$$

where $\mathbf{R}$ is the channel matrix, $\mathbf{x} = \mathbf{vec}$(ideal image), and $\mathbf{y} = \mathbf{vec}$(distorted image), and the operator $\mathbf{vec}$ simply stacks all elements of a matrix in a vector. Fig. 7.5, Fig. 7.6 and Fig. 7.7 show $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{R}$, respectively. Note that target is the letter "H". For an ideal channel, matrix $\mathbf{R}$ is diagonal, meaning that each beamlet is mapped exactly to the corresponding photo-detector. However, in the presence of turbulence, each beamlet leaks to neighboring photo-detectors as well and gives rise to IPI. As we can see from Fig. 7.7, under turbulent atmosphere conditions, $\mathbf{R}$ is no longer diagonal and received image is severely distorted.

If we have perfect knowledge of $\mathbf{R}$ matrix, we can multiply the received image vector by $\mathbf{R}^{-1}$ and invert the effects of channel. Fig. 7.8 shows restored image obtained using this inverse filtering. This approach is very similar to post-coding method used in MIMO communications, where Cholesky decomposition of channel transfer matrix is used to pre-code and post-code the transmitted and received signal vectors and hence, provide parallel channels in space. This approach can also be interpreted as classical de-convolution and inverse filtering. In other words, it is the same as case where PSF is known and can be de-convolved from the distorted observation to provide us with the original image. In the presence of noise, however, inverse filtering may cause noise enhancement and Wiener Filtering should be used which minimizes mean square error of estimation.

In the absence of such channel information (knowledge of PSF), Blind De-convolution (BD) techniques are required to estimate both PSF and original image from distorted observation.
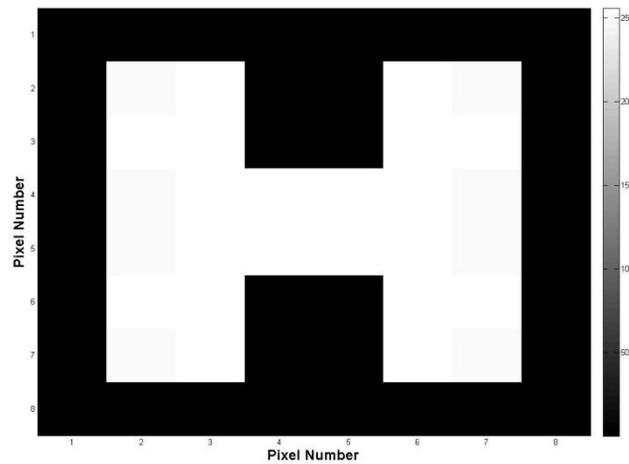
Fig. 7.5: Ideal image of letter "H".
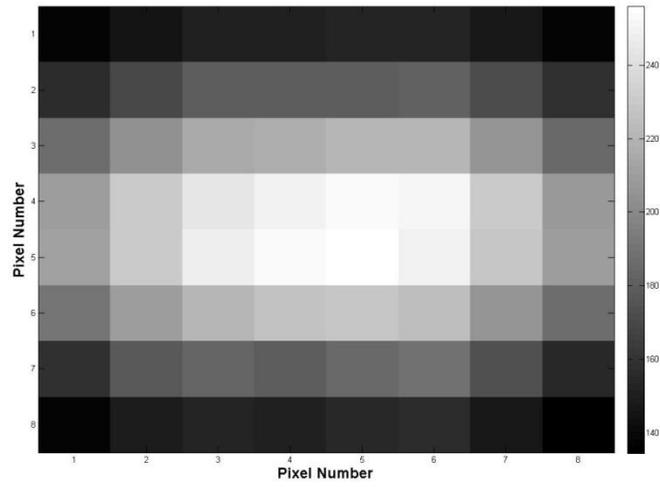
Fig. 7.6: Distorted image of letter "H".
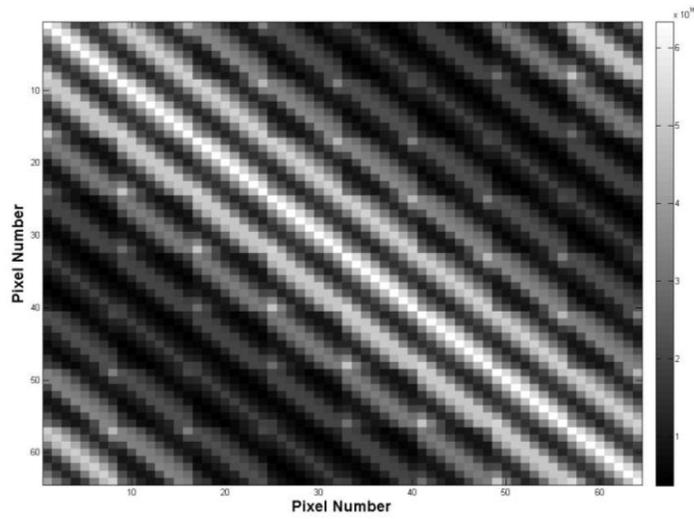
Fig. 7.7: Covariance Matrix of Channel.
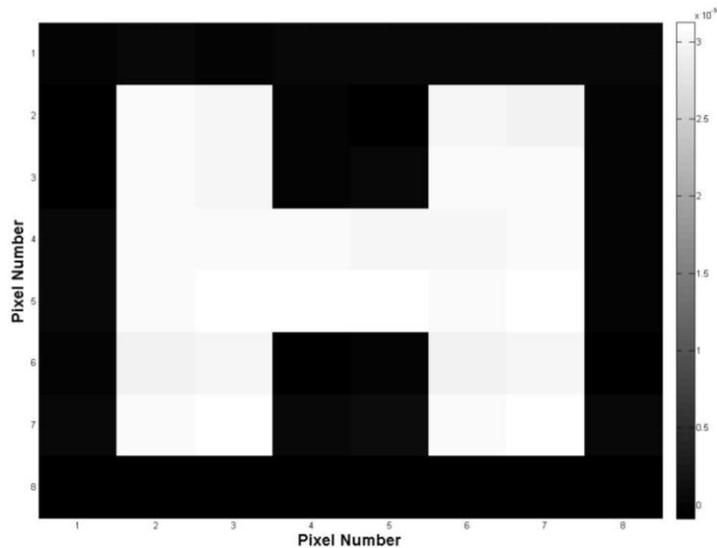
Fig. 7.8: Restored Image Using Inverse Filtering.

7.4.2 Adaptive Deconvolution

In this section, we are interested in estimating a two-dimensional target from the distorted received image. In reality, we do not have information about the target and scintillation characteristics of atmosphere change slowly with time. As a result, we have to resort to some form of blind deconvolution. Here, we present a simple outline of an adaptive deconvolution system, with the assumptions that the target is known, and the channel is constant over the time images being taken. This is similar to adaptive equalization process in RF communications systems, where a learning sequence, which is known to both transmitter and receiver, is sent through the channel and the received signal is used to estimate channel coefficients.

The undistorted target image can be represented, for example, by a 10×10 matrix $\mathbf{P}$ with each pixel/element representing the reflected intensity from a spatial coordinate (*x*, *y*) of the target. The effect of scintillation may be represented by a 100×100 matrix $\mathbf{H}$, which has the diagonal elements representing the fading experienced by individual pixels and the off-diagonal elements representing the leakage from one of the pixels to the others. The elements of $\mathbf{H}$ can be modeled as log-normal random variables with covariance matrix $\mathbf{R}_H$ denoting the covariance matrix of normal variables. This is the full correlation matrix representing the joint covariance between the elements of $\mathbf{H}$. Assuming that the correlation between the received intensities of two photo-detector elements is dependent only on their separation distance, and similar on the transmitter side, a Kronecker product model can be assumed to obtain the full correlation matrix. Therefore, $\mathbf{R}_H = \mathbf{R}_{Tx} \otimes \mathbf{R}_{Rx}$, where the transmit and receive side correlation matrices $\mathbf{R}_{Tx}$ and $\mathbf{R}_{Rx}$ can be obtained by experiments or analytical means.

With $\mathbf{R}_H$ given, $\mathbf{H}$ can be synthesized from a matrix $\mathbf{G}$ containing i.i.d. Gaussian elements as $H = \text{unvec}\left(\exp\left(2\,\mathbf{R}_H^{1/2}\text{vec}(\mathbf{G})\right)\right)$, where the $\text{vec}(\bullet)$ and $\text{unvec}(\bullet)$ operators indicate stacking of the columns of a matrix to form a column vector, and the reverse operation, respectively, and the square root indicates a matrix square root. $\mathbf{H}$ also has to be

99

properly normalized so that it neither attenuates nor amplifies the received signals, on the average. This can be done by forcing the expectation of the Froebenius norm of **H** to 1.

The resulting receiving image at the photodetectors can be written as,

$$\mathbf{V} = \text{unvec}(\mathbf{H} \cdot \text{vec}(\mathbf{P}) + \mathbf{n}) \tag{7.3}$$

where **n** represents the additive noise, modeled as a Gaussian random vector $N(\mathbf{0}, \Sigma_n)$.

Let $\mathbf{A}_k$ be the linear mean square (LMS) estimator at the kth instant in the channel tracking phase. Then the mean square error (MSE) at the kth instant is,

$$MSE_k = E\left[|\boldsymbol{\varepsilon}_k|^2\right] = E\left[|\mathbf{A}_k \text{vec}(\mathbf{V}) - \text{vec}(\mathbf{P})|^2\right] \tag{7.4}$$

For adaptive equalization, the expectation operator is ignored when taking the gradient of the MSE with respect to the elements of the LMSE estimator. As such, the gradient during K-th interval is:

$$
\begin{aligned}
\mathbf{G}_k &= \frac{\partial}{\partial \mathbf{A}_k}\left(|\mathbf{A}_k \text{vec}(\mathbf{V}) - \text{vec}(\mathbf{P})|^2\right) \\
&= \frac{\partial}{\partial \mathbf{A}_k}\left[(\mathbf{A}_k \text{vec}(\mathbf{V}) - \text{vec}(\mathbf{P}))^T (\mathbf{A}_k \text{vec}(\mathbf{V}) - \text{vec}(\mathbf{P}))\right] \\
&= 2(\mathbf{A}_k \text{vec}(\mathbf{V}) - \text{vec}(\mathbf{P}))\text{vec}^T(\mathbf{V}) \\
&= 2\boldsymbol{\varepsilon}_k \text{vec}^T(\mathbf{V})
\end{aligned}
\tag{7.5}
$$

Therefore, the estimator can be updated according to the recursion,

$$A_{k+1} = \mathbf{A}_k - \frac{1}{2} \cdot \mu \mathbf{G}_k = \mathbf{A}_k - \mu \cdot \boldsymbol{\varepsilon}_k \text{vec}^T(\mathbf{V}) \tag{7.6}$$

where μ is a step parameter suitably chosen to ensure convergence.

With the given assumptions, and the adaptive update of the LMS estimator as explained above, the results obtained for an arbitrary image P are shown in Fig. 7. (a) through (d).
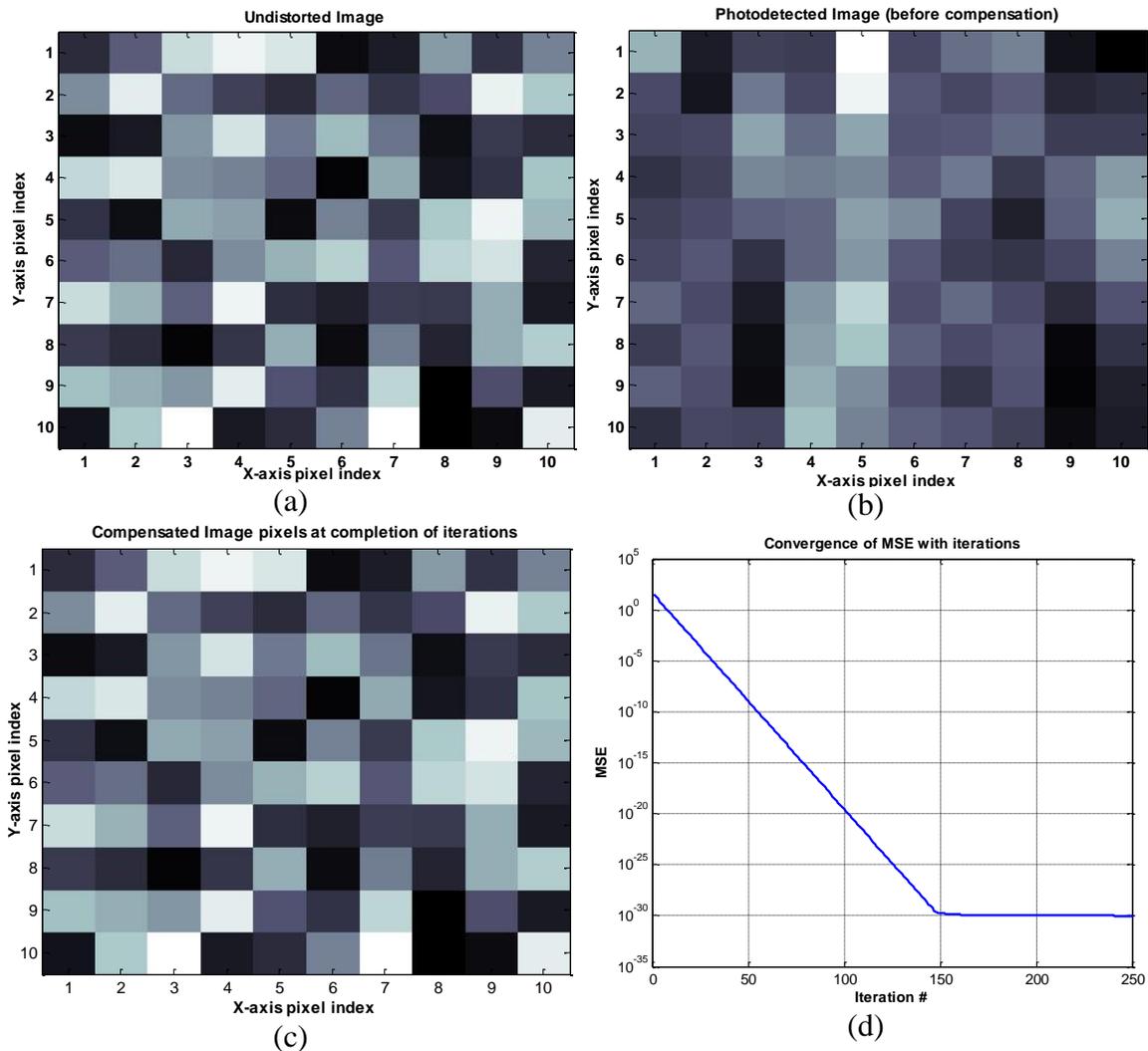
Fig. 7.9: (a) Original Image Pixels, arbitrarily chosen and known to receiver, with channel kept fixed but unknown to receiver; (b) Image pixels as detected on the photo-detector array; (c) Compensated image pixels after convergence of the adaptive deconvolution; (d) Convergence properties of the MSE with iterations.

It is noted that the estimator coefficients converge suitably within 100~150 iterations with reasonable MSE. However, the simplistic assumptions must be kept in mind, and realistic modifications must be made to the reconstruction algorithm for robustness.

7.4.3   Blind Deconvolution by Simulated Annealing

The problem of imaging in turbulent atmosphere is complicated by the fact that the imaging system does not have any information about either the true image of the object or the distorting PSF. The adaptive de-convolution algorithm is applicable only if one component of convolution is known, or can be estimated through a different mechanism up to a certain degree of accuracy. This leads us to resort to a blind mechanism which requires explicit knowledge of neither component of the convolution, but requires some information, and some

constraints have to be met. In section 6.6, we reviewed mathematical formulation of BD and showed that this formulation is applicable in both coherent and incoherent imaging systems.

A 2-D blind de-convolution problem can be formulated as:

$$g(x,y) = f(x,y) * h(x,y) + n(x,y) \tag{7.7}$$

where the true image is $f(x,y)$, the blurring PSF is $h(x,y)$ and $n(x,y)$ is the additive noise. The problem is to simultaneously estimate components $f(x,y)$ and $h(x,y)$ from $g(x,y)$. In general, it is only possible to solve the blind de-convolution problem if $f(x,y)$ and $h(x,y)$ are both of finite size.

One of the popular non-parametric blind de-convolution techniques is the Simulated Annealing (SA) algorithm [58], which is a Monte-Carlo global minimization technique that tries to minimize a given cost function. It has the following features:

1. The supports of the image and the blurring point spread function are known, i.e. $f(x,y)$ and $h(x,y)$ have finite supports $S_f(x,y)$ and $S_h(x,y)$ which are both known, beyond which $f(x,y)$ and $h(x,y)$ are zero or can be assumed very close to zero. The support refers to the smallest rectangle within which the true object is contained. Therefore, the imaging must be performed such that the object is entirely encompassed by a uniformly black, grey or white background.
2. The image and PSF pixels are non-negative.
3. Both image and PSF pixels are guessed pseudo-randomly at the beginning of the iterations, and updated simultaneously during the iterations.
4. A cost function is calculated during each iteration, which is:

$$Q(f,h,g) = \left(\sum_{x,y} (f * h - g)^2 \right) \Big/ \left(\sum_{x,y} g^2 \right) \tag{7.8}$$

Random pixel perturbations resulting in decrease of the cost function are accepted, while perturbations that increase can be accepted based on a random probability.
5. The general requirement for blind de-convolution that neither the image nor the PSF is reducible to a convolution of two or more functions holds, i.e., $f(x,y) \neq a(x,y) * b(x,y)$, for any $a(x,y)$ or $b(x,y)$.

A flow-chart of the algorithm is given in Fig. 10. The support of the true image and the PSF are known to be $M_1 \times N_1$ and $M_2 \times N_2$, respectively.

```
┌─────────────────────────────────────────────────┐
│ Inputs:                                          │
│ Blurred Image [(M₁+ M₂ -1)x(N₁+ N₂ -1)]: g       │
│ Initial Image estimate [M₁xN₁]: f = rand(0,1)    │
│ Initial PSF estimate [M₂xN₂]: h = rand(0,1)      │
│ Iteration parameters: Cycles, Scans, Nc = 1, Ns = 1 │
└─────────────────────────────────────────────────┘
```

Inputs:
Blurred Image $[(M_1 + M_2 -1) \times (N_1 + N_2 -1)]$: g
Initial Image estimate $[M_1 x N_1]$: f = rand(0,1)
Initial PSF estimate $[M_2 x N_2]$: h = rand(0,1)
Iteration parameters: Cycles, Scans, Nc = 1, Ns = 1

Normalize g so that En{g} = En{f*h}
Temperature parameter, $T = 0.01 * Q(f, h, g) / 0.8$

Nc<=Cycles?   → No → Outputs:
f, h contains de-convolved image and PSF

Yes

Ns = 1, T = 0.8 * T, $\alpha = 100\sqrt{T}$
Normalize f & h, so that En{f} = En{h}
Nc = Nc + 1

No / Yes

Perturbation $[M_1 x N_1]$ = rand(-0.5, 0.5)
f_temp = f + α * Perturbation
Set negative pixels in f_temp to 0.
Compute change in cost:
ΔQ = Q(f_temp, h, g) – Q(f, h, g)

If ΔQ>0 AND exp(-ΔQ/T)<rand(0,1): f = f,
Else f = f_temp.

Repeat step by interchanging h with f.

Ns = Ns + 1

Energy,

$$\text{En}\{p(x,y)\} = \sum_{\forall(x,y)} |p(x,y)|^2$$

Cost function,

$$Q(f,h,g) = \text{En}(f*h - g)/\text{En}(g)$$

Fig. 7.10: Restored Image Using Inverse Filtering.

To demonstrate the functionality of the SA-BD algorithm, we take a known image, $f(x, y)$, of the letter H, contained in a 6×6 pixel rectangle, and convolve it with a known 5×5 PSF, $h(x, y)$, which decays exponentially from the center point. These are shown in Fig. 7.1 (a,b).

The resulting convolution of the image and the PSF is degraded by additive noise $n(x, y)$, where the ratio of the energy of the noise to the energy of the convolved image is assumed to be 10^{-4}. The resulting noisy received image $g(x, y)$ is shown in Fig. 7.1(c). The image is similar to the case when moderate turbulence is present.

SA-BD de-convolution was performed on $g(x, y)$ with number of cycles and scans set to 100 and 500, respectively. The estimates of the image and the PSF are given in Fig. 7.1(a, b). It is apparent that the blind de-convolution algorithm was successful in separating the image and PSF. The resulting normalized squared error for the image/PSF is defined by;

$$Normalized\ Squared\ Error = \frac{\sum_{\forall(x,y)} (f - f_{est})^2}{\sum_{\forall(x,y)} f^2} \tag{7.9}$$

The error, in this case, was 0.1453 for the image of letter H and 0.0489 for the blurring PSF. The convergence of the algorithm is seen in Fig. 7.12 (a, b). The temperature parameter, T, is reduced logarithmically throughout the cycles, which controls the perturbation of the pixel values. The cost function diminishes to about 10^{-5}, after about 80 cycles.

True Image: Letter H (a)

PSF (b)

Received Noisy Image (c)

Estimated PSF (d)
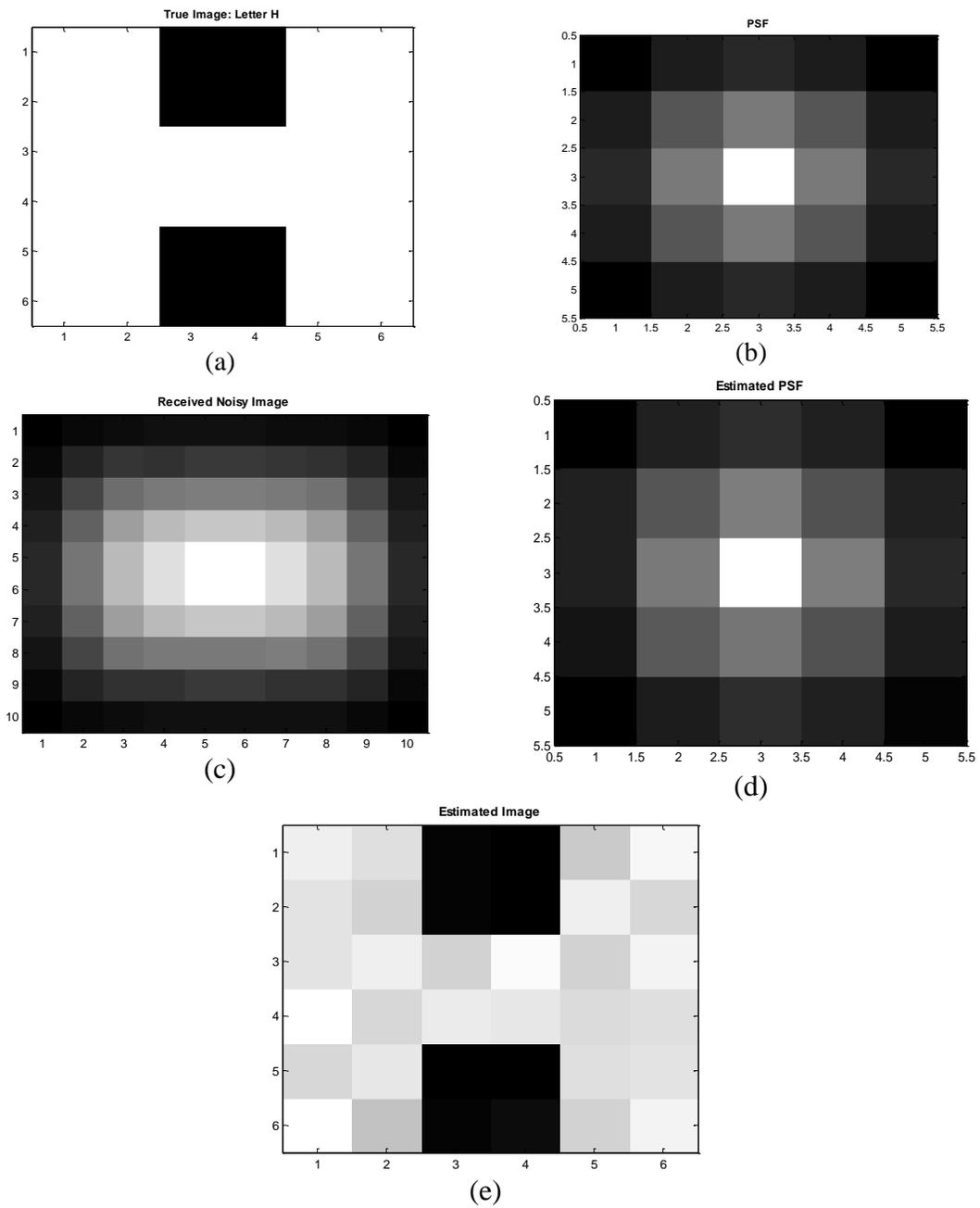
Estimated Image (e)

Fig. 7.11: (a) Original image to be detected; (b) Blurring Point-Spread Function; (c) Convolution of Image and PSF degraded by Additive Noise; (d) Estimated PSF, (e) Estimated Image
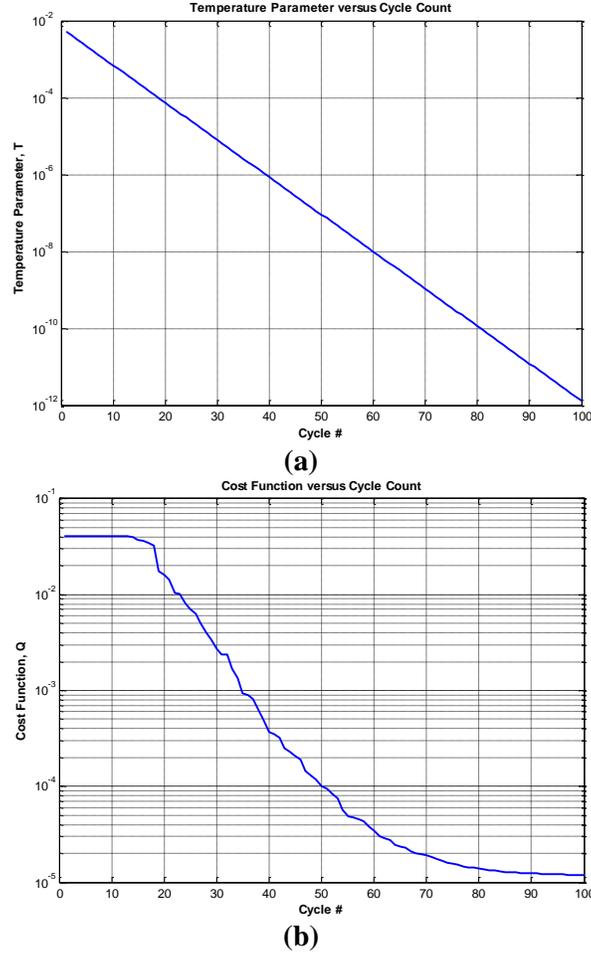
Fig. 7.12 (a) Change of Temperature Parameter with Cycles; (b) Convergence of
the Cost Function.

Two performance measures, namely Mean Square Error (MSE) and Signal to Noise Ratio Improvement (SNRI) are used to quantify this improvement. MSE is defined as

$$MSE(\hat{f}) = 100 \frac{\sum [a\hat{f}(x,y) - f(x,y)]^2}{\sum f(x,y)^2} \tag{7.10}$$

where recovered image is shown by $\hat{f}(x,y)$ and original image is shown by $f(x,y)$. Furthermore, to account for scaling of images, parameter a is used to minimize the MSE and is given by:

$$a = \frac{\sum f(x,y)\hat{f}(x,y)}{\sum \hat{f}(x,y)^2} \tag{7.11}$$

Fig. 7.13, Fig. 7.14 and Fig. 7.15 present three different cases of image recovery under weak, moderate, and strong turbulence conditions and investigate the convergence and

performance of SABD in each of these situations. Fig. 7.13 through Fig. 7.15 show distorted image, recovered image, and MSE percentage at each iteration. We see that SABD operates properly at all turbulence levels. However, note that convergence of SABD depends on the initial guess of image, which is random. As a result, the recovered image would not be the same every time. Here, we chose to show the best results obtained using this algorithm.

(a)

(b)

(c)
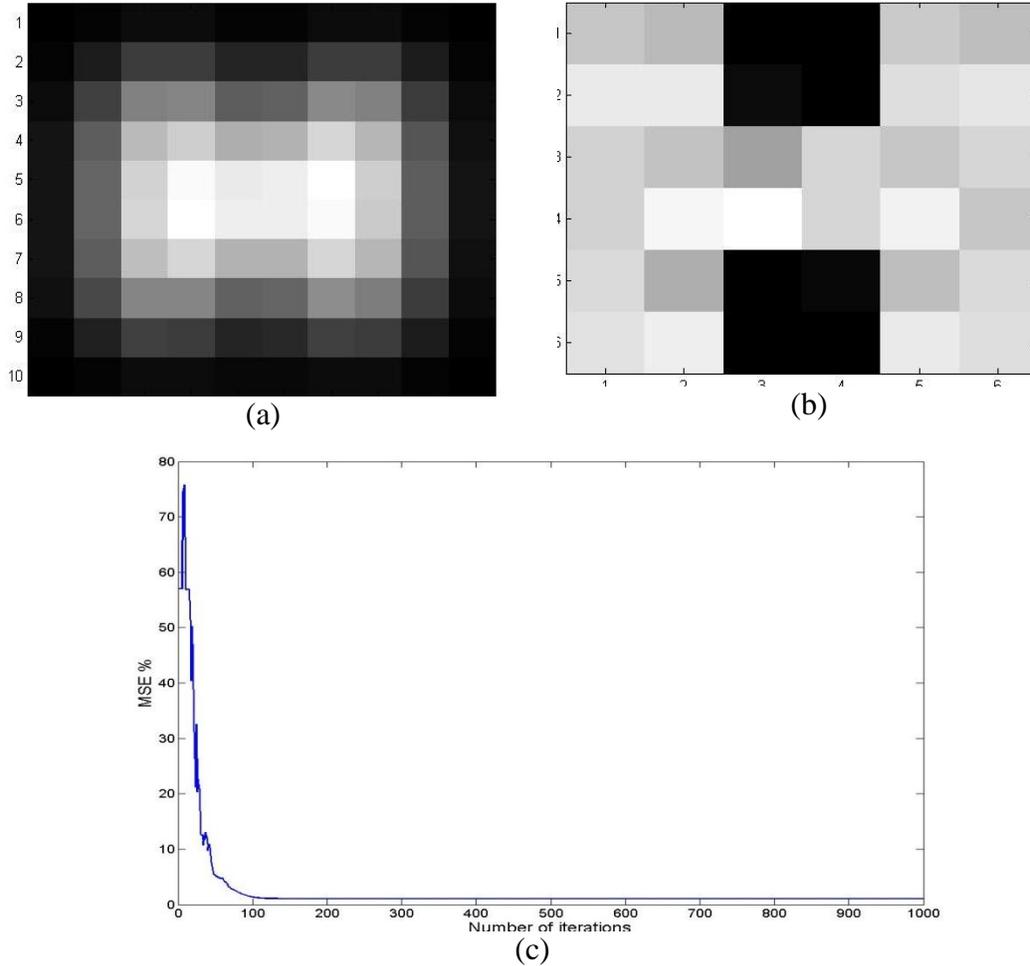
Fig. 7.13: (a) Distorted Image, (b) Recovered Image, (c) MSE % per iteration.
Weak turbulence condition.
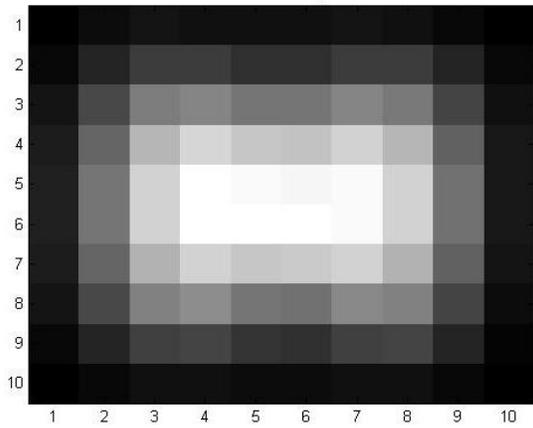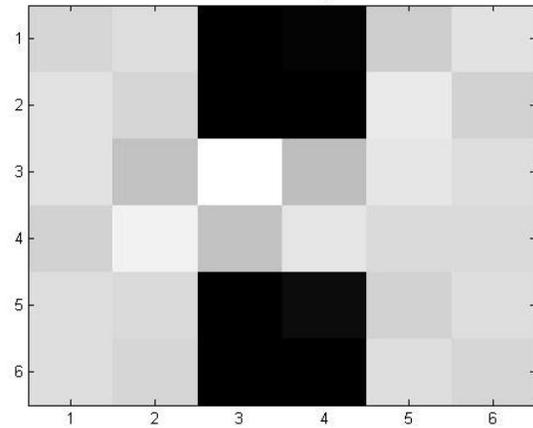
(a)   (b)
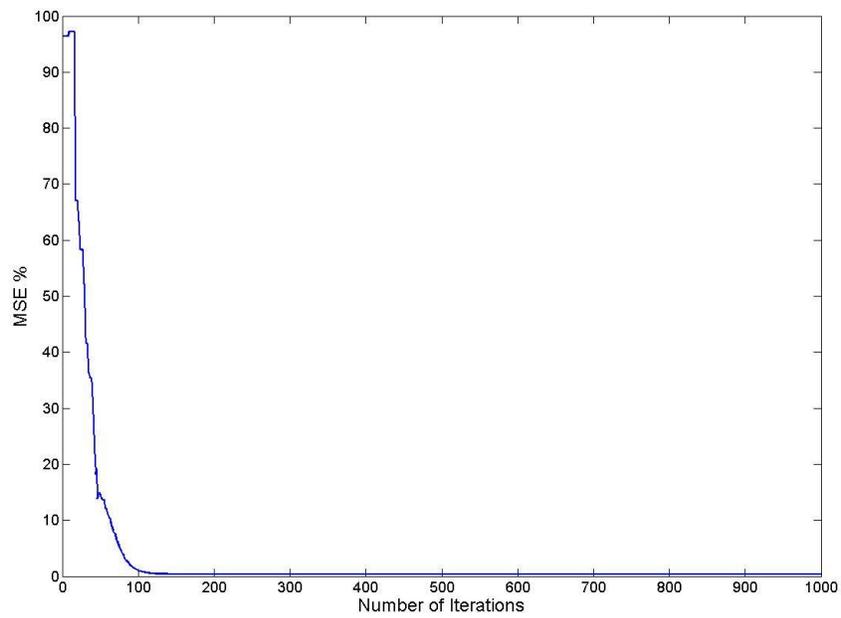
(c)

Fig. 7.14: (a) Distorted Image, (b) Recovered Image, (c) MSE % per iteration. Moderate turbulence condition
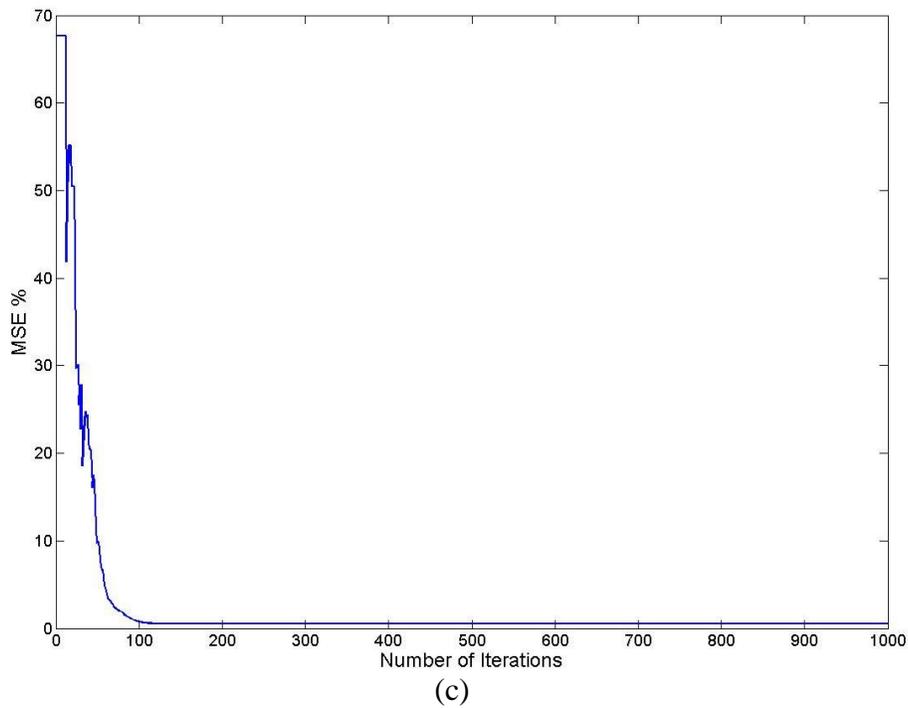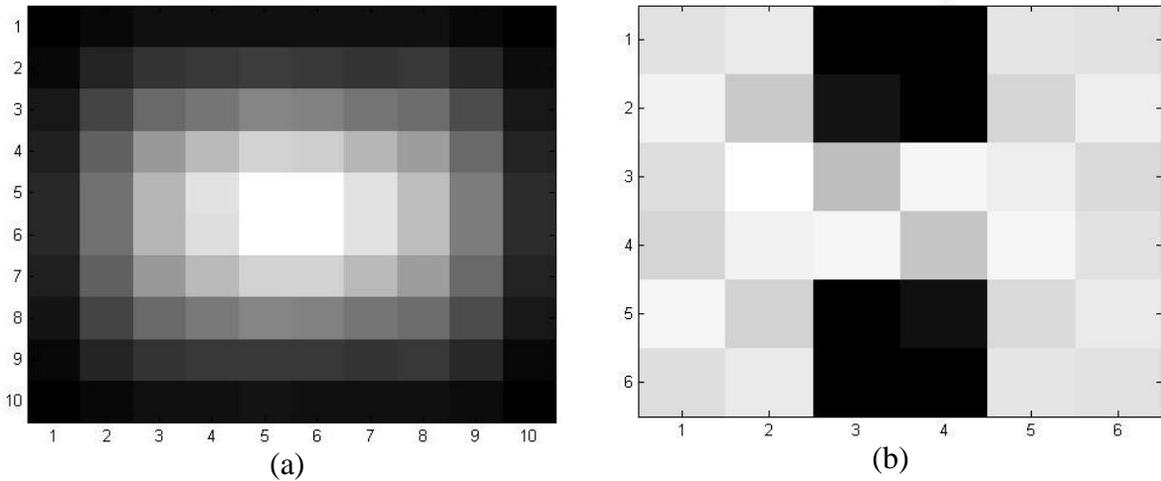
Fig. 7.15: (a) Distorted Image, (b) Recovered Image, (c) MSE % per iteration. Strong turbulence condition.

One can observe that MSE percentage decreases fast. Furthermore, in the examples above, SABD provides 13.7 dB, 18.8 dB, and 18.6 dB SNR improvements for weak, moderate, and strong turbulence conditions, respectively. Again, note that SABD performance highly depends on the initial guess of original image and the above figures are examples of SABD performance with a "good" initial state.

7.4.4 Non-Negativity and Support-constrained Recursive Inverse Filtering (NAS-RIF)

The NAS-RIF algorithm is developed to solve a version of BD problem, defined as follows:

Given a gray-scale image **g(x,y)** *degraded by a linear shift invariant PSF* **h(x,y)**, *find a reliable estimate of the true image* **f(x,y)**, *given partial or no information about the PSF and true image.*

Non-negativity and support information can extrapolate high spatial frequency components lost by a band-limiting distortion. Hence, non-negativity and support constraints are used as partial information to define a cost-function. The true image is estimated subject to minimization of this error criterion or cost function, which is defined as:

$$J = \sum_{(x,y) \in D_{neg}} \hat{f}(x,y)^2 + \sum_{(x,y) \in D_{sup}} \left(\hat{f}(x,y) - L_B \right)^2 \tag{7.12}$$

where D_{sup} is the set of all pixel locations of $\hat{f}(x,y)$ within the region of support and D_{neg} is the set of all pixel location within D_{sup}, for which $\hat{f}(x,y)$ is negative. Furthermore, $\hat{f}(x,y)$ is the estimated true image and L_B is the average pixel value of the background. This cost-function is proved to be convex, and thus having a global minima [59]. To recover the true image, $f(x,y)$, the received blurred image is filtered by a variable FIR filter $u(x,y)$ to produce an estimate of the image $\hat{f}(x,y)$. This estimate is passed through a nonlinear filter, which projects the estimate onto the convex set corresponding to the non-negativity and support constraints of the true image. The difference between this projected image $\hat{f}_{NL}(x,y)$ and $\hat{f}(x,y)$, i.e. the cost function or the error signal, is used to update the variable filter coefficients. Block diagram of the proposed algorithm is shown in Fig. 7.16 [59].
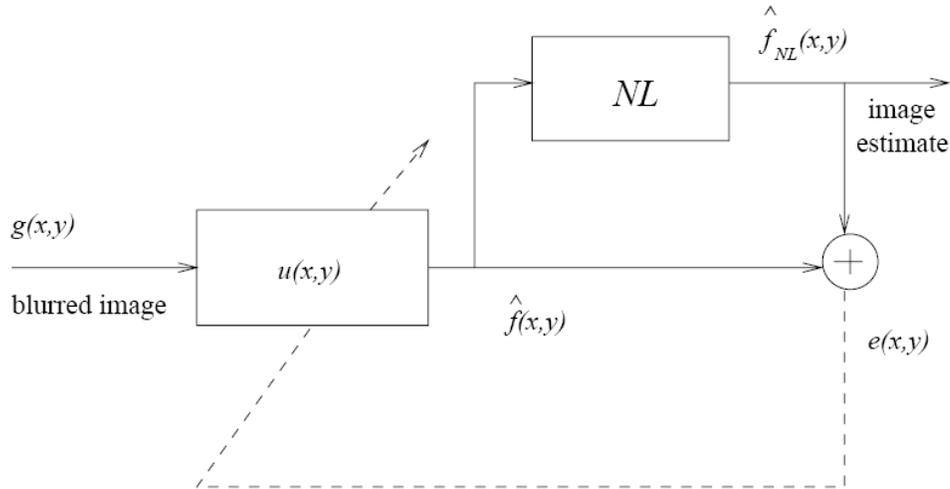
Fig. 7.16: Block Diagram of NAS-RIF algorithm

In order to minimize the convex cost-function of NAS-RIF, Conjugate-Gradient optimization is used, which finds a direction that minimizes the cost function at each iteration. The summary of NAS-RIF algorithm that uses conjugate gradient algorithm for optimization is shown in Table 7.2 [59].

110

Table 7.2: Summary of the NAS-RIF method using the conjugate gradient algorithm for optimization.

• Set initial conditions (k=0) $u_k^T = [u_k(1,1) \quad \ldots \quad u_k\big((N_{xu}+1)/2,(N_{yu}+1)/2\big) \quad \ldots \quad u_k(N_{xu},N_{yu})]$ $\quad = [0,\ldots,1,\ldots,0]$ $\quad\quad$ set $\delta > 0$.
• At iteration (k): k = 0, 1, 2, … $\quad$ 1) If $J(u_k) \le \delta$, stop. $\quad$ 2) If $k = 0$, $d_k = -\nabla J(u_k)$. Otherwise, $d_k = -\nabla J(u_k) + \beta_{k-1}d_{k-1}$. $\quad$ 3) Perform a line minimization to find t_k such that $\quad\quad\quad J(u_k + t_k d_k) \le J(u_k + td_k)$ for $t \in \Re$. $\quad$ 4) $u_{k+1} = u_k + t_k d_k$ where $u_k \in R^{N_{xu} \times N_{yu}}$ is the current estimate of the inverse filter , $\big[\nabla J(u_k)\big]_{i,j} = \dfrac{\partial J(u_k)}{\partial u(i,j)} = 2 \displaystyle\sum_{(x,y)\in D_{neg}} \hat{f}(x,y)g(x-i+1,y-j+1)$ $\quad\quad\quad + 2 \displaystyle\sum_{(x,y)\in D_{sup}}[\hat{f}(x,y)-L_B]g(x-i+1,y-j+1)$, and $\beta_k = \dfrac{\big\langle \nabla J(u_{k+1}) - \nabla J(u_k), \nabla J(u_{k+1})\big\rangle}{\big\|\nabla J(u_k)\big\|^2}$.

7.4.5   Zero Sheet Separation Method

Though it seems impossible at first glance, it has been shown that components of a convolution in a space of dimensions more than unity can be recovered simply from the convolution itself. This argument is based on the fact that the roots of the spectrum (either Fourier-transform or Z-transform) of an N-dimensional image (2-D in our case) lie on a $(2N - 2)$-dimensional hyper surface in a $2N$-dimensional space (N-dimensional complex) [60]. This hyper surface is called the "Zero Sheet" of an image. The spectrum of a composite image is the product of the original image and PSF spectra and hence its roots are given by the union of PSF and original image zero sheets. Suppose that the distorted image is given by:

$$g(\mathbf{x}) = f(\mathbf{x}) * h(\mathbf{x}) \tag{7.13}$$

This translates to the following in transform domain (either Fourier or Z transform):

$$G(u,v) = F(u,v)H(u,v) \tag{7.14}$$

Hence, roots of $G(u,v)$ are the union of $F(u,v)$ and $h(u,v)$ roots. Now, we elaborate the concept of zero sheets in more details. We know that the transform domain representation of a pixilated image, $f(m,n)$ is given by:

$$F(u,v) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f(m,n)u^m v^n \qquad (7.15)$$

For a fixed u, (7.15) can be written as:

$$F(u,v) = \sum_{n=0}^{N-1} a_n(u)v^n \qquad (7.15)$$

where

$$a_n(u) = \sum_{m=0}^{M-1} f(m,n)u^m \qquad (7.16)$$

Hence, $F(u,v)$ can be written as a 1-D polynomial whose coefficients depend on the data $f(m,n)$ and the complex variable u [61]. For any value of u, this polynomial has $M-1$ roots that depend on u. Zero sheets are formed by starting from an initial $u = u_0$, successively increasing u and evaluating the roots of the polynomial in (7.15). The increments in u are small enough that the displacement of each point zero can be clearly determined from one increment to the next. The same steps can be followed by the role of u and v being reversed.

If one follows these steps for $G(u,v)$, zero sheets of $F(u,v)$ and $h(u,v)$, being necessarily continuous, are unambiguously and separately mapped out. In the presence of noise, however, we have:

$$G(u,v) = F(u,v)H(u,v) + N(u,v) \qquad (7.17)$$

As a result, $G(u,v)$ can no longer be factorized and disjoint hyper-surfaces merge into a single hyper-surface. Hence it would not be possible to discriminate between roots of PSF and original image. Fig. 7.17 (a) through (d) show the zero sheets of original image, PSF, distorted image , and distorted noisy image, respectively. It is clear that points in Fig. 7.17(c) are the union of points in Fig. 7.17(a) and Fig. 7.17(b). Furthermore, due to continuity of zero sheets, it is theoretically possible to find the components of a composite image. However, when noise is added to distorted image, disjoint, continuous and closed traces of the PSF and original image zeros merge and can not be discriminated any more.

Fig. 7.1 (a) through (d) show original image, PSF, distorted image, and distorted noisy image, respectively. We observe that the received image is severely distorted. However, from the above argument, it is still possible to recover the original image. In the presence of noise, however, zero sheets patterns have changed, significantly. This is due to the fact that the coefficients of polynomials have been perturbed and hence, zeros are displaced. Note that, the Signal-to-Noise Ratio (SNR) level is quit high (40 dB) and the effect of noise is not clear from Fig. 7.1 (d).

In the subsequent section, we investigate the possibility of restoring the original polynomial of composite image spectra from the perturbed polynomial.
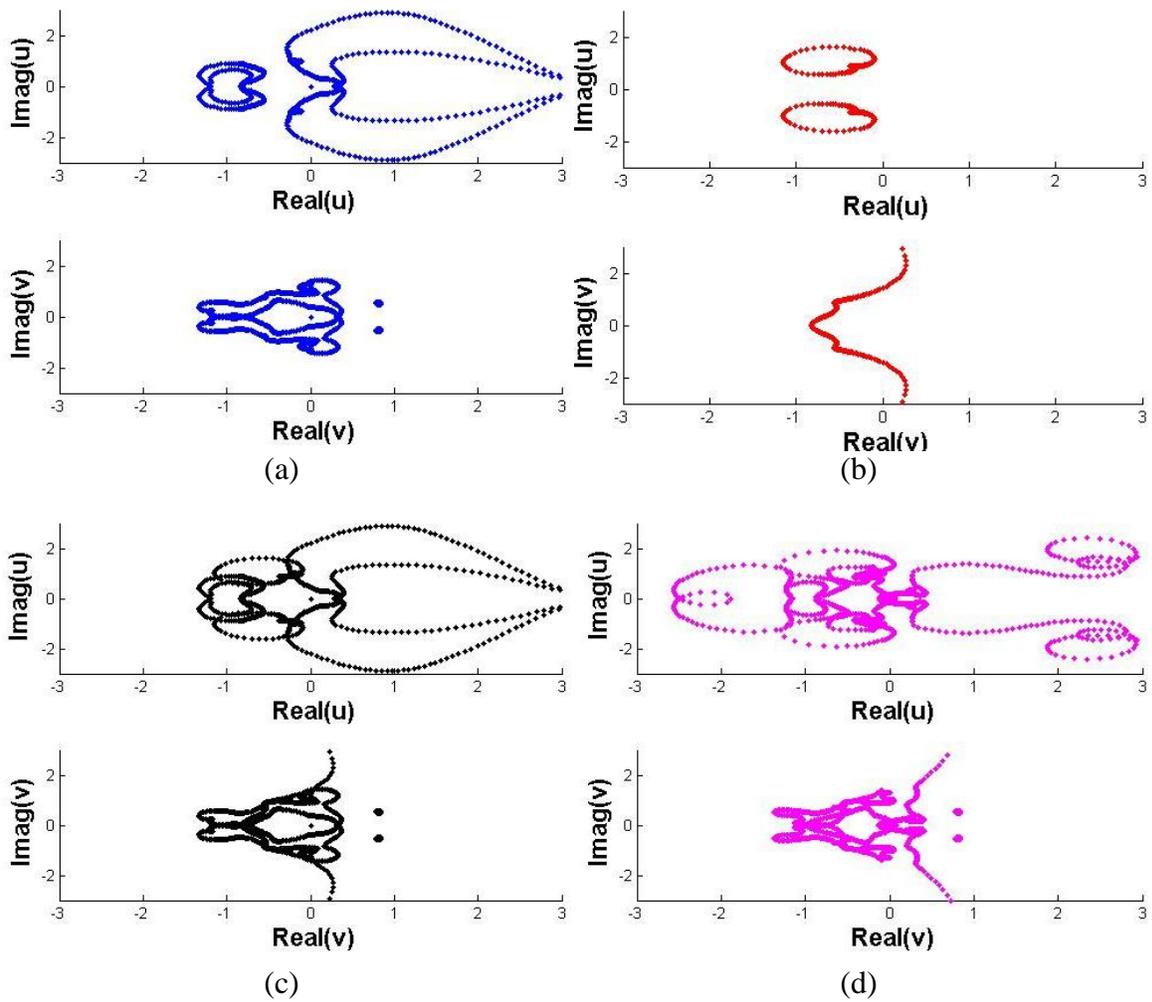
Fig. 7.17: (a) Zero sheets of original image; (b) Zero sheets of PSF; (c) Zero sheets of distorted image; (d) Zero sheets of distorted & noisy image
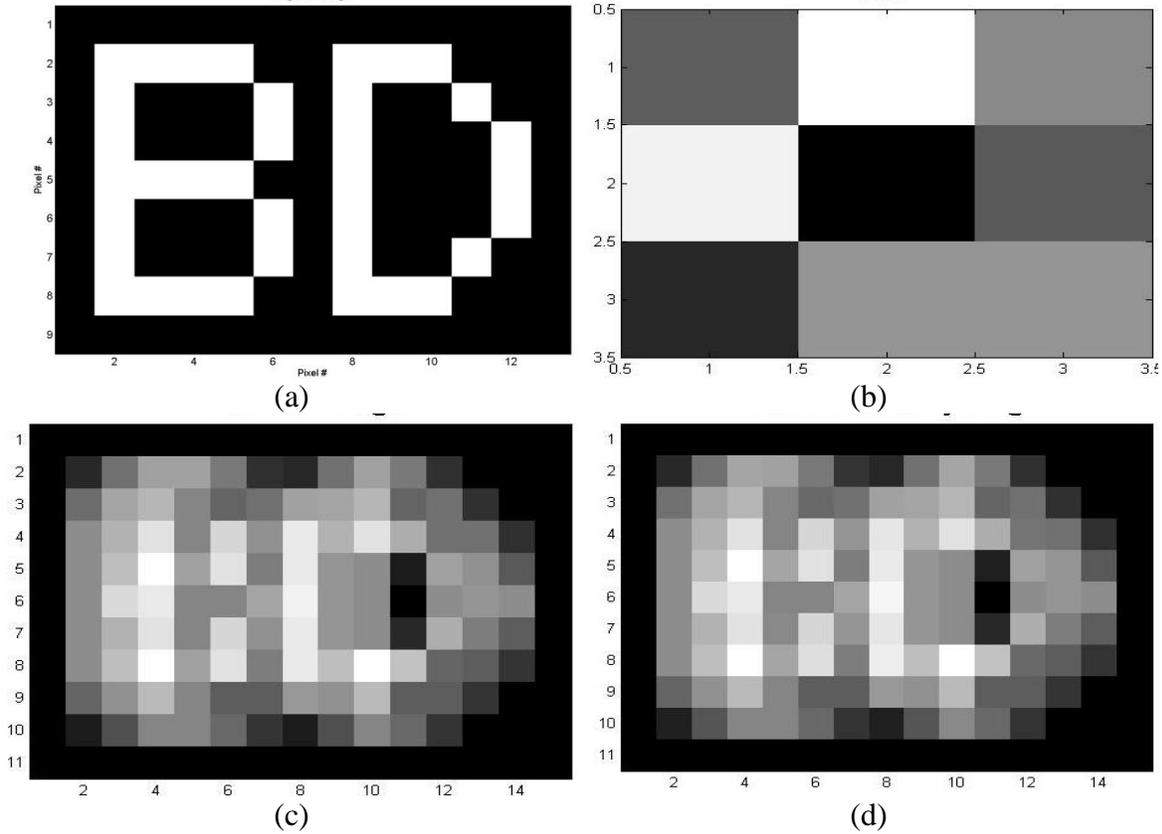
Fig. 7.18: (a) Image; (b) PSF; (c) Distorted image; (d) Distorted & noisy image

7.4.6 Blind Deconvolution Based on Approximate Factorization of Bivariate Polynomials (*AFBP*)

As in (7.17), the *z*-transform of the blurred image, $G(z_1, z_2)$, is a bivariate polynomial in variables z_1 and z_2; hence, the problem of deconvolution becomes the problem of factorizing the polynomial in z_1 and z_2. Unfortunately, in practice, the coefficients of the polynomial are not perfect, and are corrupted either by additive noise in the imaging system or due to finite precision error in the detectors. Therefore, an algorithm is required to be devised that allows for approximate factorization of this bivariate polynomial. The references [62][63] provide the mathematical bases for factorizing polynomials, and the criterion for approximate factorization is to find the "nearest" factorizable polynomial having a total degree less than or equal to the given polynomial, in the least square sense.

Letting $G(z_1, z_2) \in \mathbf{Q(i)}(z_1, z_2)$ be an irreducible polynomial over the set of complex numbers $\mathbf{C}$, where irreducibility is caused by perturbations on the coefficients of *G*. Then the algorithm presented in [62] finds $G^{[\min]}$, a factorizable polynomial over $\mathbf{C}$ with $\deg(G^{[\min]}) \le \deg(G)$, such that the normalized error $\|G - G^{[\min]}\|_2 / \|G\|_2$ is minimized.

The algorithm is based on Ruppert's criterion of irreducibility of polynomials. If $G(z_1, z_2) \in \mathbf{C}(z_1, z_2)$ with bi-degree (M, N), i.e., $\deg_{z_1} G = M$, $\deg_{z_2} G = M$, then Ruppert's criterion states that *f* is absolutely irreducible if and only if the equation

114

$$\frac{\partial}{\partial z_2}\left(\frac{F}{G}\right)=\frac{\partial}{\partial z_1}\left(\frac{H}{G}\right) \qquad (7.18)$$

has no nonzero solution $F,H \in C[z_1,z_2]$ with $\deg F \le (M-1,N)$, $\deg H \le (M,N-2)$.

Since (7.18) is linear over C, it gives a linear system for the coefficients of F and H, whose coefficient matrix is called the Ruppert matrix $R(G)$. The matrix $R(G)$ is full rank if and only if G is absolutely irreducible. Furthermore, a basis of the solution space can be found from the Ruppert matrix which could be used, in conjunction with a multivariate GCD finding algorithm, to find out the approximate factorization of G.

Letting $G(z_1,z_2) \in Q(i)(z_1,z_2)$ be an irreducible polynomial over the set of complex numbers C, where irreducibility is caused by perturbations on the coefficients of G. Then the algorithm presented in [62] finds $G^{[\min]}$, a factorizable polynomial over C with $\deg\left(G^{[\min]}\right) \le \deg(G)$, such that the normalized error $\left\|G - G^{[\min]}\right\|_2 / \|G\|_2$ is minimized.

The following polynomial, presented as a benchmark first in [62],

$$\begin{aligned}
G(z_1,z_2) &= z_1^5 + z_1^3 z_2^2 - z_1^2 z_2 + 7z_1^2 + z_1^4 z_2 + z_1^2 z_2^3 - 2z_1 z_2^2 \\
&\quad +7z_1 z_2 + 2z_1^3 z_2 + 2z_1 z_2^3 - 2z_2^2 + 15z_2 - z_1^3 - 7 + 0.2z_1 \\
&= z_1^5 z_2^5 \left(\begin{array}{l} z_2^{-5} + z_1^{-2}z_2^{-3} - z_1^{-3}z_2^{-4} + 7z_1^{-3}z_2^{-5} + z_1^{-1}z_2^{-4} + z_1^{-3}z_2^{-2} - 2z_1^{-4}z_2^{-3} + 7z_1^{-4}z_2^{-4} \\ +2z_1^{-2}z_2^{-4} + 2z_1^{-4}z_2^{-2} - 2z_1^{-5}z_2^{-3} + 15z_1^{-5}z_2^{-4} - z_1^{-2}z_2^{-5} - 7z_1^{-5}z_2^{-5} + 0.2z_1^{-4}z_2^{-5} \end{array} \right)
\end{aligned} \qquad (7.19)$$

is formed by perturbing the product of the following two polynomials by $0.2z_1$,

$$F(z_1,z_2) = z_1^2 + z_1 z_2 + 2z_2 - 1 = z_1^2 z_2^3 \left(z_2^{-3} + z_1^{-1}z_2^{-2} + 2z_1^{-2}z_2^{-2} - z_1^{-2}z_2^{-3} \right) \qquad (7.20)$$

$$H(z_1,z_2) = z_1^3 + z_1 z_2^2 - z_2 + 7 = z_1^3 z_2^2 \left(z_2^{-2} + z_1^{-2} - z_1^{-3}z_2^{-1} + 7z_1^{-3}z_2^{-2} \right) \qquad (7.21)$$

The algorithm described in [62] performs the approximate bivariate factorization on the polynomial with perturbed coefficients, and yields,

$$\begin{aligned}
\tilde{F}(z_1,z_2) &= 0.000341124534691889555 \cdot z_2^3 - 0.294509433380196140 \cdot z_2^2 z_1 \\
&\quad -0.000041378972244730443 \cdot z_1^2 z_2 - 0.296197230468099148 \cdot z_1^3 \\
&\quad +0.0097128963538054020 2 \cdot z_2^2 - 0.0142725188110806103 \cdot z_2 z_1 \\
&\quad +0.0129003734808113332 \cdot z_1^2 + 0.291409870076061728 \cdot z_2 \\
&\quad +0.0109341660279600808 \cdot z_1 - 2.038754149 \\
&= z_1^3 z_2^3 \cdot \left(\begin{array}{l} 0.000341124534691889555 \cdot z_1^{-3} - 0.294509433380196140 \cdot z_1^{-2}z_2^{-1} \\ -0.000041378972244730443 \cdot z_1^{-1}z_2^{-2} - 0.296197230468099148 \cdot z_2^{-3} \\ +0.0097128963538054020 2 \cdot z_1^{-3}z_2^{-1} - 0.0142725188110806103 \cdot z_1^{-2}z_2^{-2} \\ +0.0129003734808113332 \cdot z_1^{-1}z_2^{-3} + 0.291409870076061728 \cdot z_1^{-3}z_2^{-2} \\ +0.0109341660279600808 \cdot z_1^{-2}z_2^{-3} - 2.038754149 \cdot z_1^{-3}z_2^{-3} \end{array} \right)
\end{aligned} \qquad (7.22)$$

with the constant multiplier, $c = -6.676951989$, so that,

$$G^{[min]}(z_1, z_2) = c \cdot \tilde{F}(z_1, z_2) \cdot \tilde{H}(z_1, z_2) \tag{7.23}$$

The normalized error of the approximate factorization is,

$$\left\| G - G^{[min]} \right\|_2 \Big/ \left\| G \right\|_2 = 0.01084227541 \tag{7.24}$$

By least squares, the scaling factors of the individual functions are found out, and the reconstructed image and blurring functions are given by,

$$F_{factored}(z_1, z_2) = -3.431574113610501 \cdot \tilde{F}(z_1, z_2) \tag{7.25}$$

$$H_{factored}(z_1, z_2) = 1.941112961 \cdot \tilde{H}(z_1, z_2) \tag{7.26}$$

The normalized error for each of these are found out similarly as in (7.24),

$$err_1 = 0.010845341474554, \quad err_2 = 0.011829151573117 \tag{7.27}$$

The normalized error for each of these are found out similarly as in (7.24). These results are summarized in Fig. 7.19(a-j), as 2-D images.

The above example only deals with a large additive noise value in one pixel. Further investigations are required to find out the performance of the algorithm when additive noise perturbs all the pixels independently, or in a correlated manner.

The approximate multivariate factorization algorithm may have several advantages over conventional blind deconvolution mechanisms which are based on Bussgang methods or the simulated annealing method described in the previous monthly report. The last two methods are iterative and require tremendous computing power. While it is not still clear how much complexity the AFBP (Approximate Factorization Bivariate Polynomials) algorithm would impose on current processors, but it is a completely different approach to deal with the blind deconvolution problem. We intend to investigate the computational complexity of the algorithm, as well as its performance, and compare them with previously presented approaches, in our future works.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)
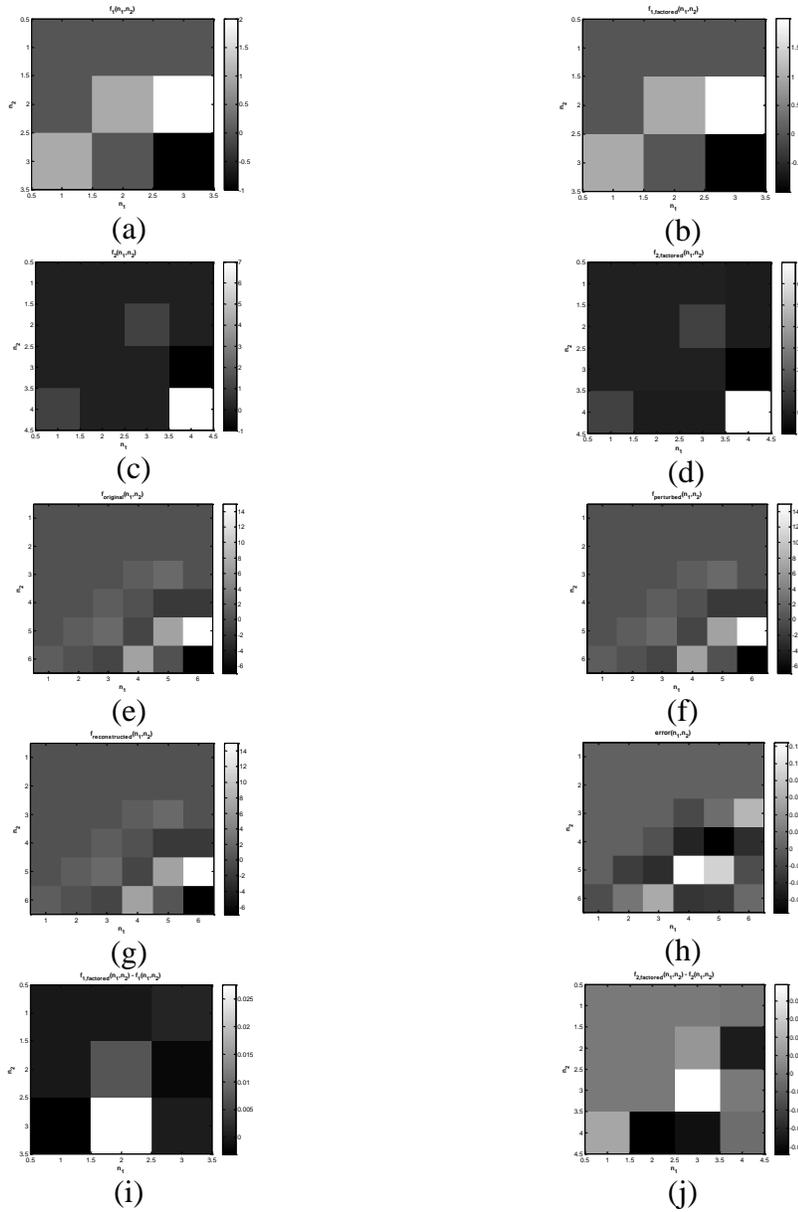
Fig. 7.19: Summary of polynomial factorization results: (a) Original Image; (b) Rescaled Factored Image, $f_{,\text{factored}}$; (c) Blur function, h; (d) Rescaled blur function, $h_{,\text{factore.}}$; (e) Convolved Image, g_{original}; (f) Convolved image with perturbation, $g_{\text{perturbed}}$; (g) Reconstructed convolved image from factored polynomials, $f_{\text{reconstructed}}$; (h) Error between reconstructed and perturbed original image; (i) Error between original image and factored image; (j) Error between original blur function and factored blur function

We present a more detailed set of results with the 8×8 image of the letter 'H', which we have been using as a standard. It is subjected to blurring by seven sample PSF's, namely:
(a) PSF1: Horizontal blur, physically corresponding to horizontal motion,
(b) PSF2: Vertical blur, corresponding to vertical motion of camera or object,
(c) PSF3: Diagonal blur, corresponding to diagonal motion of camera or object,

(d) PSF4: PSF due to turbulence with wind speed of 21 m/s and A = 1.7×10^{-14},
(e) PSF5: PSF due to turbulence with wind speed of 21 m/s and A = 1.7×10^{-13},
(f) PSF6: PSF due to turbulence with wind speed of 21 m/s and A = 1.7×10^{-12},
(g) PSF7: PSF due to turbulence with wind speed of 21 m/s and A = 5×10^{-12}.

The first three PSF's (PSF1-3) represent the simplest types of blurring possible, and hence can be used as benchmarks to qualify the severity of degradations imposed on the true image by the PSF's due to turbulence (PSF4-7), and the performance of the proposed algorithm.

We present performances for two different cases:

(a) *Noiseless environment:* This is the ideal case where additive noise is absent. However, due to quantization in the digital electronics, the true coefficients are mapped to a set of discrete values, and the factorization or deconvolution still has to be performed with approximate coefficients. The deconvolution results presented in this case are obtained with the default floating point precision of Matlab/Maple.

(b) *Noisy environment:* In this case, i.i.d.Gaussian noise is added to the blurred image pixels, and the measure of the noise is given by the blurred SNR or BSNR, defined as,

$$BSNR = \frac{\sum_{x,y} |g(x, y)|^2}{E\left\{\sum_{x,y} |n(x, y)|^2\right\}} = \frac{E_g}{MN \cdot \sigma_n^2} \tag{7.28}$$

where $g(x, y)$ is the blurred image pixel values at locations (x, y), E_g is the energy contained in the blurred image, $n(x, y)$ is the additive Gaussian noise at each pixel, with mean 0 and variance σ_n^2, and MN is the total number of pixels, where M is the number of rows in the photo-detector array, and N is the number of columns.

Fig. 7. through Fig. 7.2 present the blind deconvolution results obtained with the standard 'H' image with the PSF's mentioned in the last section, for the noiseless case. Fig. 7.28 through Fig. 7.34 are for the noisy case when the BSNR is set at 150dB, according to the definition given in the last section.

Fig. 7.20. True Image

(a) Blurring PSF

(b) Blurred Image

(c) Reconstructed PSF
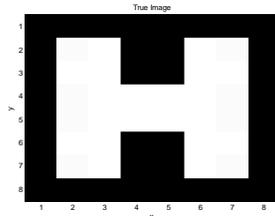
(d) Reconstructed Image
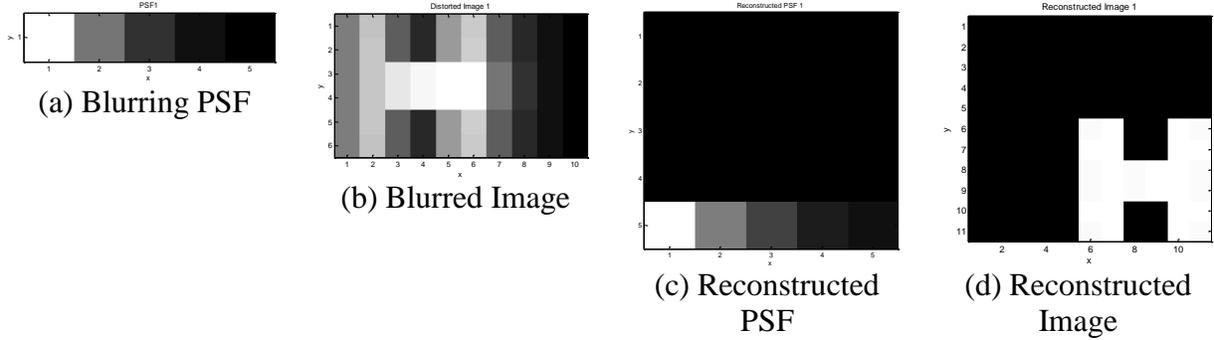
Fig. 7.21. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF1, noiseless case

(a) Blurring PSF

(b) Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
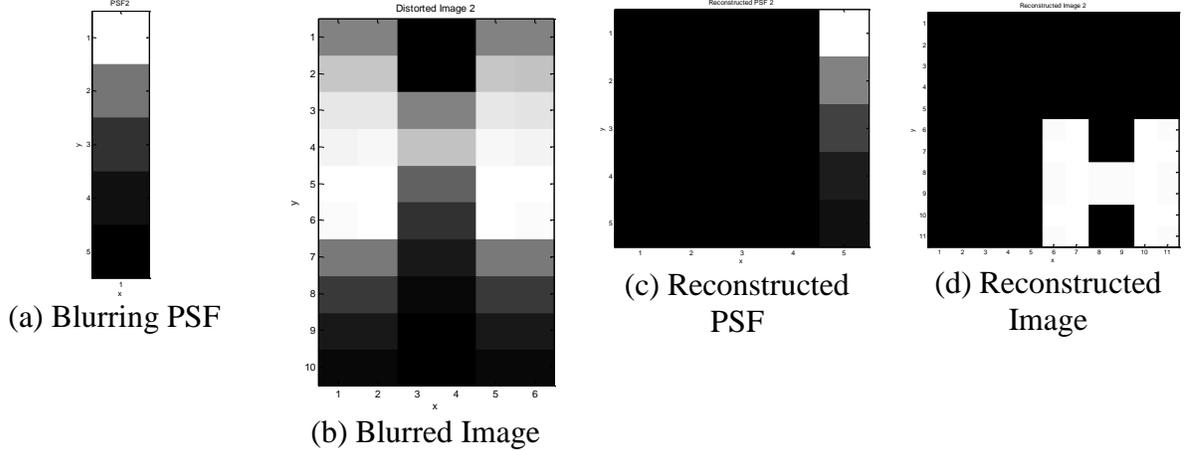
Fig. 7.22. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF2, noiseless case

(a) Blurring PSF

(b) Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
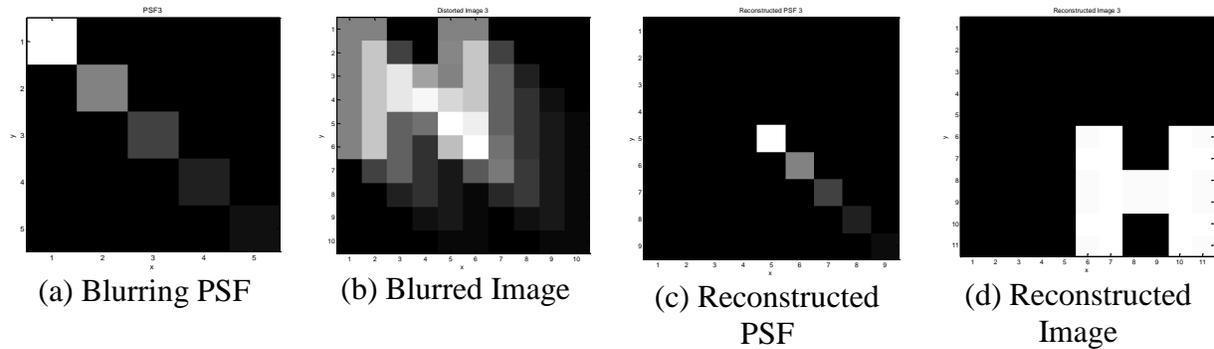
Fig. 7.23. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF3, noiseless case

(a) Blurring PSF

(b) Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
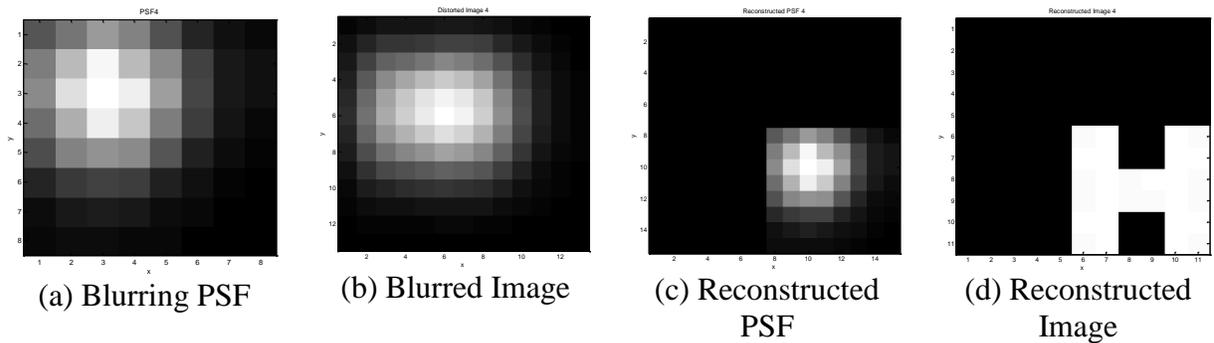
Fig. 7.24. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF4, noiseless case

119

(a) Blurring PSF (b) Blurred Image (c) Reconstructed PSF (d) Reconstructed Image
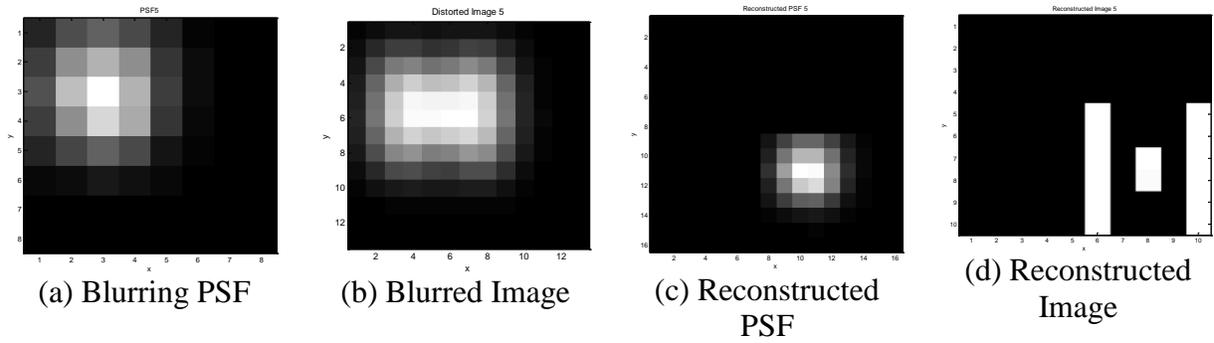
Fig. 7.25. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF5, noiseless case

(a) Blurring PSF (b) Blurred Image (c) Reconstructed PSF (d) Reconstructed Image
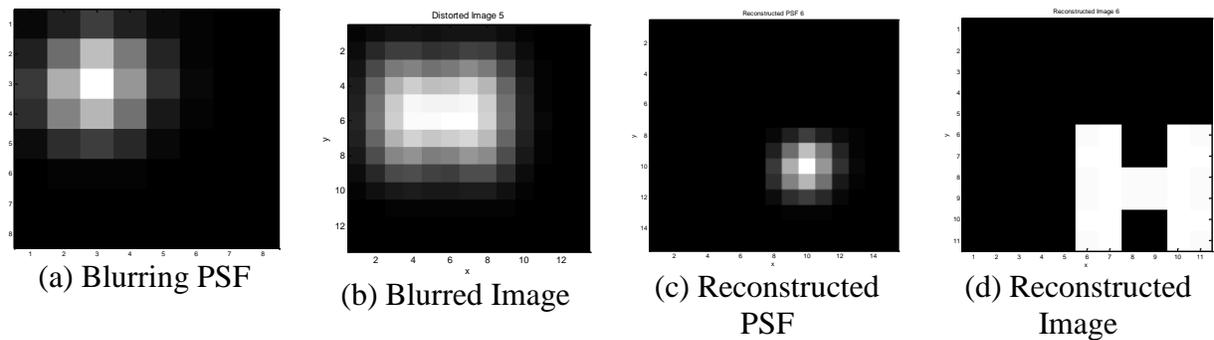
Fig. 7.26. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF6, noiseless case

(a) Blurring PSF (b) Blurred Image (c) Reconstructed PSF (d) Reconstructed Image
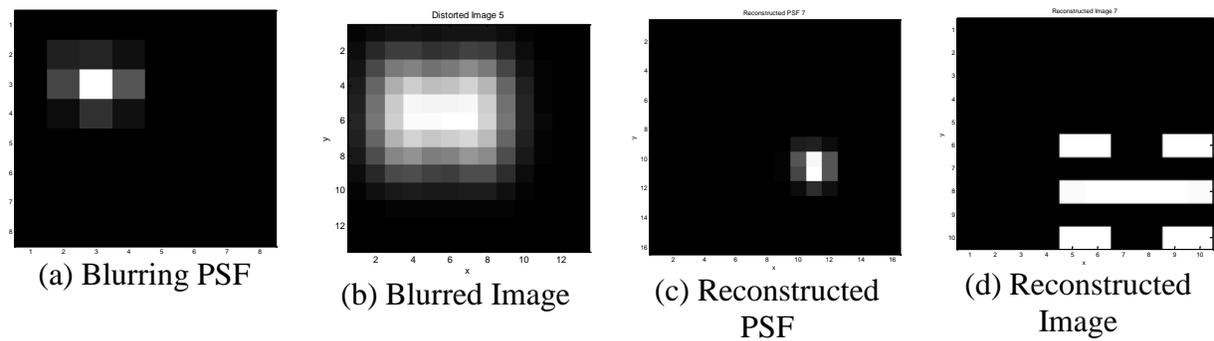
Fig. 7.27. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF7, noiseless case

(a) Blurring PSF (b) Noisy Blurred Image (c) Reconstructed PSF (d) Reconstructed Image
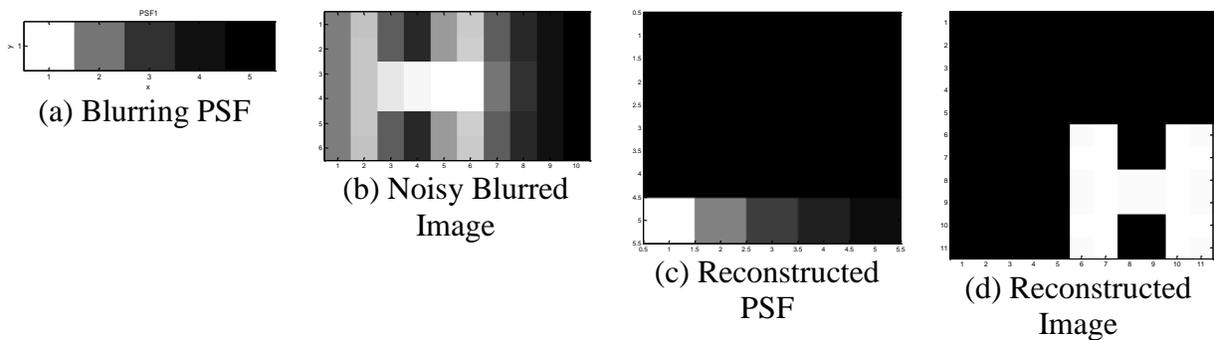
Fig. 7.28. Blurring PSF, Noisy Blurred Image, Reconstructed PSF and image for PSF1, BSNR=150dB

(a) Blurring PSF

(b) Noisy Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
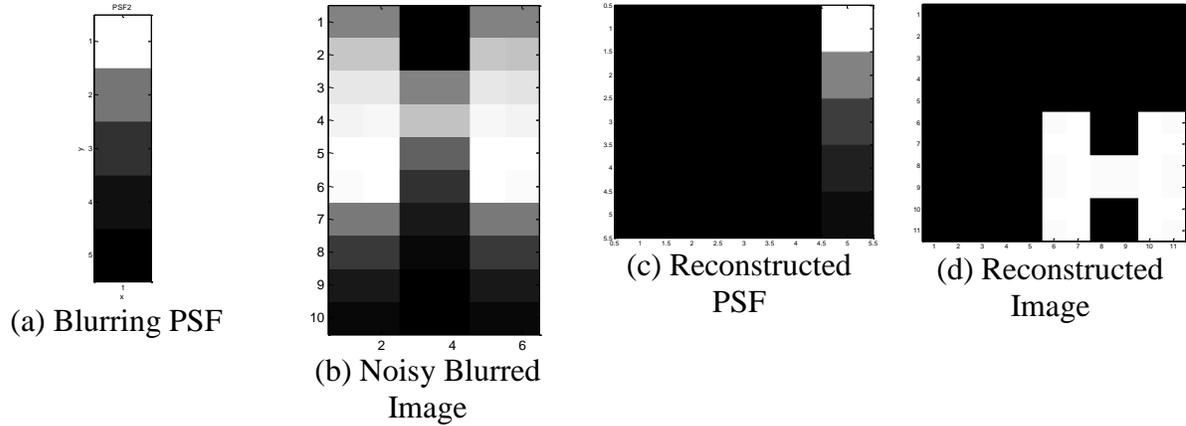
Fig. 7.29. Blurring PSF, Noisy Blurred Image, Reconstructed PSF and image for PSF2, BSNR=150dB

(a) Blurring PSF

(b) Noisy Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
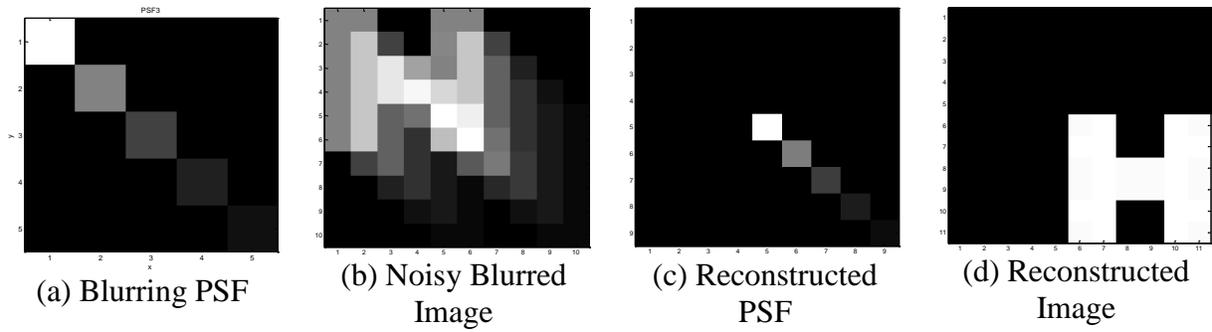
Fig. 7.30. Blurring PSF, Noisy Blurred Image, Reconstructed PSF and image for PSF3, BSNR=150dB

(a) Blurring PSF

(b) Noisy Blurred Image

(c) Reconstructed PSF

(d) Reconstructed Image
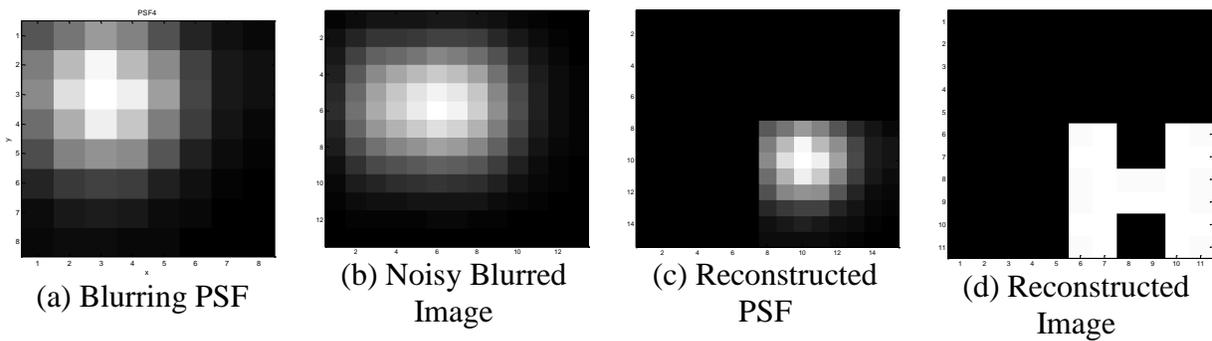
Fig. 7.31. Blurring PSF, Noisy Blurred Image, Reconstructed PSF and image for PSF4, BSNR=150dB

(a) Blurring PSF   (b) Noisy Blurred Image   (c) Reconstructed PSF   (d) Reconstructed Image
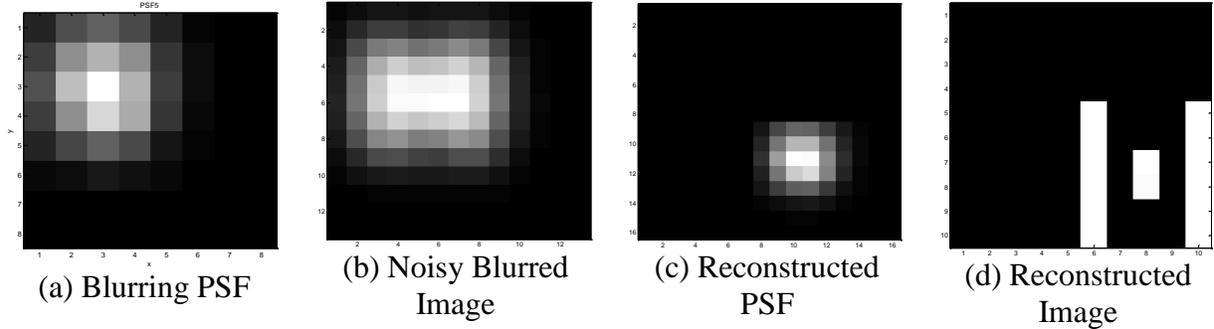
Fig. 7.32. Blurring PSF, Noisy Blurred Image, Reconstructed PSF and image for PSF5, BSNR=150dB

(a) Blurring PSF   (b) Noisy Blurred Image   (c) Reconstructed PSF   (d) Reconstructed Image
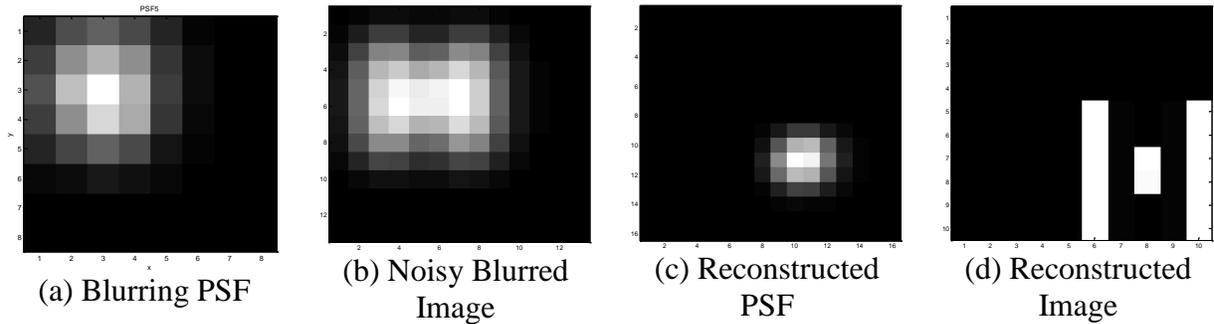
Fig. 7.33. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF6, BSNR=150dB

(a) Blurring PSF   (b) Blurred Image   (c) Reconstructed PSF   (d) Reconstructed Image
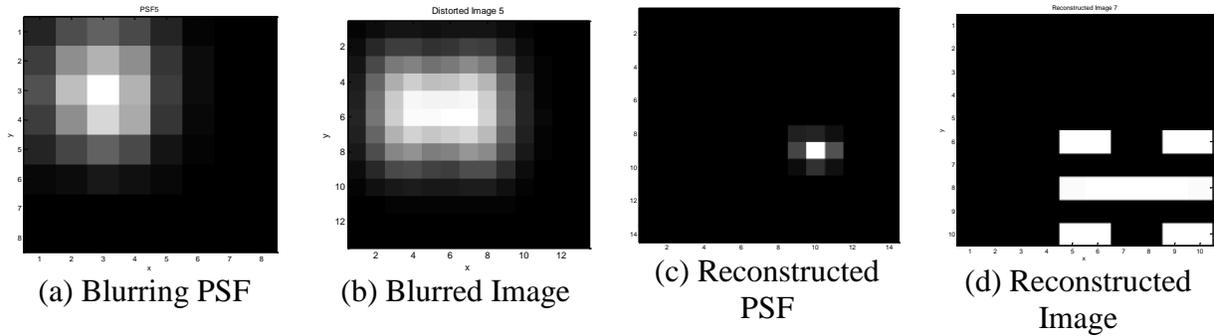
Fig. 7.34. Blurring PSF, Blurred Image, Reconstructed PSF and image for PSF7, BSNR=150dB.

The performance of the blind deconvolution algorithm can be quantified in terms of percentage MSE, defined as,

$$\%MSE = 100 \cdot \frac{\sum_{x,y}\left[a\hat{f}(x,y) - f(x,y)\right]^2}{\sum_{x,y}\left|f(x,y)\right|^2} \qquad (7.29)$$

where $\hat{f}(x, y)$ is the recovered image/PSF, $f(x, y)$ is the true image/PSF, and the scaling term a is chosen such that MSE is minimized. Another performance measure is the SNR improvement (SNR) with respect to the blurred image, which is defined as,

$$SNRI = \frac{\%MSE(g)}{\%MSE(\hat{f})} \tag{7.30}$$

where $g(x, y)$ is the blurred and possibly noisy image.

Table 7.3 summarizes the SNRI and %MSE values for the noiseless case, while the same values for BSNR values of 150dB and 140dB are tabulated in Table 7.4 and Table 7.5.

Table 7.3. Performances of the proposed Blind Deconvolution in noiseless scenario.

Noiseless Reconstruc-tion	SNRI, dB	% MSE of Reconstructed Image	% MSE of PSF	% MSE of Observed Image
PSF1	132.7666	8.2152×10^{-13}	6.3201×10^{-13}	1.3878×10^{-13}
PSF2	132.4715	6.5212×10^{-13}	$5.5456e \times 10^{-14}$	4.3025×10^{-13}
PSF3	134.3180	6.8362×10^{-13}	1.0246×10^{-11}	1.7545×10^{-12}
PSF4	85.7649	1.0620×10^{-7}	1.4386×10^{-8}	9.0872×10^{-11}
PSF5*	-	49.3455	7.5340	6.6671×10^{-5}
PSF6	68.8632	3.3075×10^{-6}	1.7165×10^{-6}	4.8206×10^{-10}
PSF7*	-	49.3583	32.1110	1.0560×10^{-6}
* indicates that the deconvolution did not yield the true image.				

Table 7.4. Performances of the proposed Blind Deconvolution algorithm for BSNR = 150dB

Noisy Reconstruction (BSNR = 150dB)	SNRI, dB	% MSE of Reconstructed Image	% MSE of PSF	% MSE of Observed Image
PSF1	101.7536	1.0374×10^{-9}	1.4394×10^{-9}	2.0619×10^{-10}
PSF2	111.0314	9.0853×10^{-11}	1.1475×10^{-10}	3.8488×10^{-11}
PSF3	123.4641	8.3216×10^{-12}	1.5512×10^{-10}	4.0366×10^{-11}
PSF4	93.7879	1.6742×10^{-8}	2.8398×10^{-9}	1.4165×10^{-11}
PSF5*	-	49.3502	7.5318	5.8235×10^{-5}
PSF6*	-	49.3713	9.4677	1.1355×10^{-8}
PSF7	25.9326	0.0192	0.0237	3.6554×10^{-6}
* indicates that the deconvolution did not yield the true image.				

Table 7.5. Performances of the proposed Blind Deconvolution algorithm for BSNR = 140dB

Noisy Reconstruction (BSNR = 140dB)	SNRI, dB	% MSE of Reconstructed Image	% MSE of PSF	% MSE of Observed Image
PSF1	112.7325	8.2801×10^{-11}	1.8797×10^{-10}	1.7016×10^{-11}
PSF2	94.9802	3.6598×10^{-9}	1.3769×10^{-8}	1.2429×10^{-9}
PSF3	105.4036	5.3242×10^{-10}	2.4992×10^{-9}	2.2239×10^{-10}
PSF4*	-	49.3003	7.7710	8.9320×10^{-3}
PSF5*	-	49.3344	7.5529	7.6154×10^{-6}
PSF6*	-	49.3697	9.4657	1.0995×10^{-8}
PSF7*	-	49.4023	32.0138	1.0041×10^{-4}
* indicates that the deconvolution did not yield the true image.				

Examining the results shows that the approximate factorization of bivariate polynomials is not always successful in deconvolving the true image from the blur, even in the noiseless case. This is due to the fact that the constraint is imposed only on the distorted received image/polynomial, and the algorithm tries to find the "nearest" factorizable polynomial, in the least square sense. Moreover, the algorithm seems to break down for lower BSNR values. Further limitations that were observed with the algorithm are as follows:

a. The true image to be recovered must be on a uniformly black or white or grayscale background, so that a 'full-field' deconvolution is possible. A 'full-field' deconvolution means that the true image support is completely within the blurred image support. The algorithm fails to achieve appropriate results if pixels values are not zero beyond the support of the blurred image. This is a limitation encountered with all non-parametric blind deconvolution algorithms [59][64]. This scenario is valid for astronomical speckle imaging, and medical imaging [59].

b. If the blurred image has more than two factors, there is an ambiguity about the true image, and some a priori information or pattern matching criterion has to be used to obtain the true image. This is a general limitation of all blind deconvolution techniques [64].

c. The algorithm is very sensitive to additive noise. As we can see even with a BSNR of 140dB, the algorithm fails to perform proper deconvolution and recovery of the true image, with the PSF's corresponding to atmospheric turbulence. However, it still performs well with the three simpler blurring functions at this BSNR value. Since the image restoration problem is ill conditioned [64], a small perturbation of the given data produces a large deviation in the solution.

To improve the performance of the algorithm, performance under the following modifications can be investigated:

a. Take multiple frames of the object to be imaged, and use overlap-and-add technique to obtain a new blurred image. Each exposure would be degraded by independent noise samples, and also possibly convolved with uncorrelated blurring functions. If the exposures are taken fast enough, i.e. within the coherence time of turbulence, the induced blur may not change significantly. The resulting polynomial representation of the image can then be given by,

$$\tilde{g}(z_1, z_2) = \frac{1}{N}\left[\sum_{i=1}^{N}\{f(z_1, z_2)g_i(z_1, z_2) + n_i(z_1, z_2)\}\right]$$

$$= \frac{1}{N}\left[f(z_1, z_2)\cdot\sum_{i=1}^{N}g_i(z_1, z_2) + \sum_{i=1}^{N}n_i(z_1, z_2)\right] \qquad (7.30)$$

The shifts can be determined from the 2-D cross-correlation peaks from one frame to the next. Furthermore, if the additive noise is i.i.d between frames, the resulting added noise will have lower variance; as a result, the BSNR, and hence the performance of the proposed deconvolution technique, may be expected to improve.

b.   In addition to the 'least-square' error constraint imposed on the distorted image, further constraints may be imposed on the recovered true image. This algorithm currently does not use any a priori information about the true image. If constraints are imposed on the deconvolved image, such as known support, non-negativity, continuity over a certain range, statistical distribution, entropy, known parametric model of either the PSF or the object, etc. [64], we may expect to see some performance improvement of this algorithm.

c.   When imaging through turbulent atmosphere, we could use some adaptive optics compensation technique to compensate for lower Zernike orders. The resulting image would then be post-processed using the factorization algorithm.

7.5   Comparison of NAS-RIF and AFBP

We investigate the performance improvements obtained using the Non-Negativity and Support constraints Recursive Inverse Filtering (NAS-RIF) algorithm [64], as well as our Approximate Factorization of Bivariate Polynomials (AFBP) method and demonstrate Mean Square Error (MSE) percentage and Signal-to-Noise Ratio Improvements (SNRI) achieved via post-processing. Note that in [64], SNRI is defined, even for the noise-free conditions, as the ratio of MSE before and after processing. Hence, it would be more appropriate to call it Signal-to-Noise plus Interference Ratio Improvements.

Fig. 7.35 (a), (b), and (c) show the simulation results for focal plane image of an *8×8* MIMO imaging system under ideal, as well as weak and moderate turbulence conditions, respectively, where object to be imaged is the letter "H". Fig. 7.3 (a), (b), and (c) show the simulation results for the photo-detected images under ideal, weak, and moderate turbulence conditions, respectively. It is clear from Fig. 7.3 that as turbulence strength increases, the photo-detected image becomes blurrier.

To recover the original image, receiver should be well-equipped with an Adaptive-Optics (AO) system. Moreover, proper post-processing techniques can be exploited to enhance the image quality. In this section, we assume a simple AO system that only corrects for the tip-tilt component of phase distortion and leaves the major burden of image recovery to the post-processing or BD module.

Fig. 7.3 (a) and (b) show the reconstructed image under weak turbulence conditions using NAS-RIF and AFBP methods, respectively. Furthermore, Fig. 7.3 (c) shows the recovered PSF using AFBP under a weak turbulence condition. Note that NAS-RIF is based on inverse filtering and only generates the inverse of PSF. Fig. 7.3 (d) and (e) show the reconstructed image under strong turbulence conditions using NAS-RIF and AFBP methods, respectively. Furthermore, Fig. 7.3 (f) shows the recovered PSF using AFBP under a strong turbulence

condition. From Fig. 7.3, one can see that under a weak turbulence condition, both NAS-RIF and AFBP have a good performance. However, under a strong turbulence condition, NAS-RIF is not as efficient as AFBP method. Nonetheless, NAS-RIF is much more robust to noise, compared to AFBP. Table 7.6 and Table 7.7 provide quantitative comparison of NAS-RIF and AFBP using SNRI and MSE percentage as a performance criterion, respectively.

From Table 7.6, it is inferred that SNRI is quite the same for NAS-RIF and AFBP under a weak turbulence condition. However, while NAS-RIF fails to provide sufficient SNRI under a strong turbulence condition, AFBP shows an even larger SNRI compared to the weak turbulence condition. Unfortunately, single frame AFBP is not robust to noise and we are trying to improve this technique by exploiting multiple frames. Note that the noise is added such that input average SNR per pixel is 70dB [64]. NAS-RIF can also operate at average SNR level of 30-40 dB, but there is no guarantee that the algorithm will converge. Table 7.7 shows the MSE percentage at the output of BD module. Again, one can observe that under noise-free conditions, AFBP is superior to NAS-RIF. Another interesting observation is that AFBP performs better under strong turbulence conditions. This is due to the fact that the two bivariate polynomials that are recovered by AFBP, namely the z-transform of PSF and original image, are of comparable degrees under strong turbulence conditions and hence are more distinguishable.

| (a) | (b) | (c) |
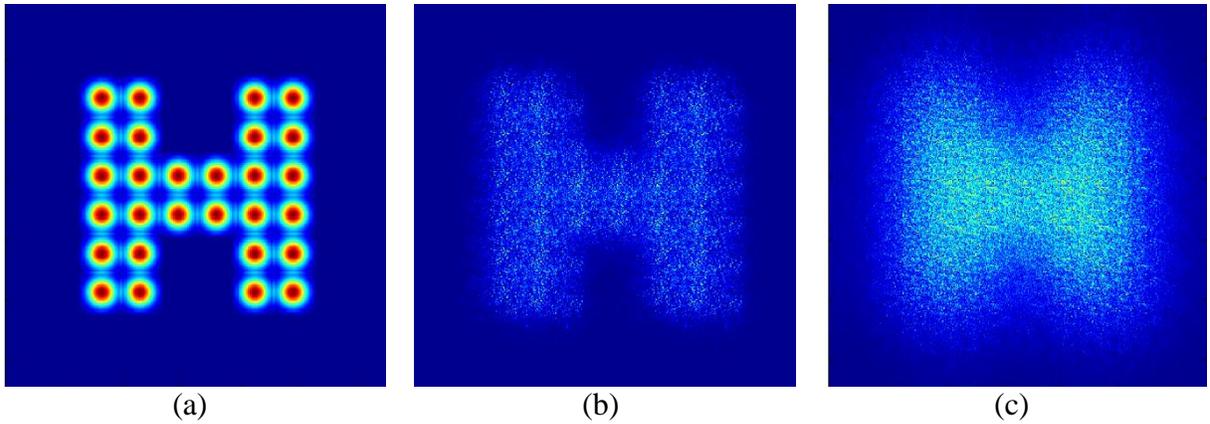
Fig. 7.35 Focal plane images of an *8×8* MIMO imaging system, (a) ideal, (b) weak turbulence, and (c) moderate turbulence.

| (a) | (b) | (c) |
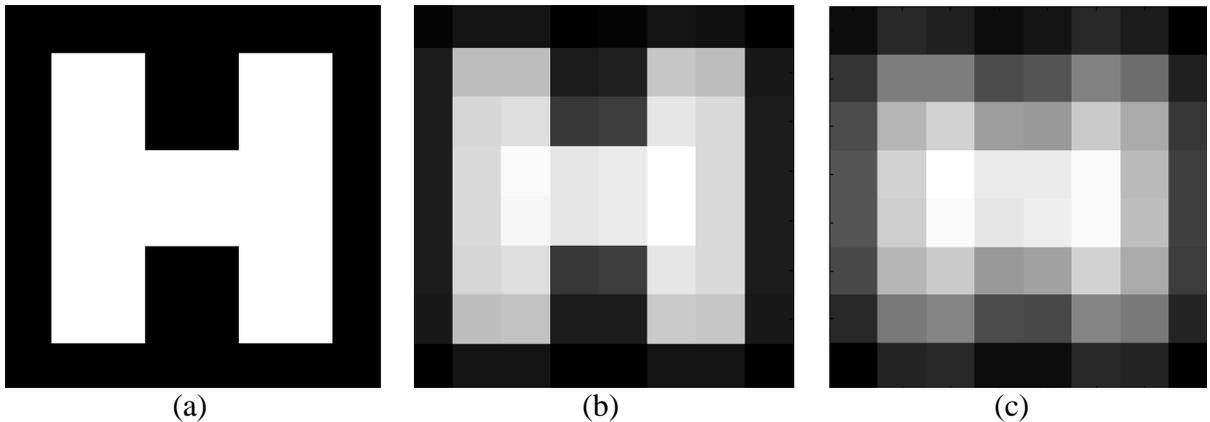
Fig. 7.36 Photo-Detected images of an *8×8* MIMO imaging system, (a) ideal, (b) weak turbulence, and (c) moderate turbulence.

Table 7.6 SNRI for NAS-RIF and AFBP algorithms under weak and strong turbulence conditions, with and without Noise.

SNRI	Weak Turbulence	Strong Turbulence	Weak Turbulence+ Noise	Strong Turbulence+ Noise
NAS-RIF	37dB	16.5 dB	36.7 dB	16.2 dB
AFBP	41 dB	77 dB	N/A	N/A

Table 7.7 MSE percentage for NAS-RIF and AFBP algorithms under weak and strong turbulence conditions, with and without Noise.

MSE %	Weak Turbulence	Strong Turbulence	Weak Turbulence+ Noise	Strong Turbulence+ Noise
NAS-RIF	0.0007 %	0.53 %	0.0007%	0.56 %
AFBP	2.7985×10^{-4}	4.8725×10^{-7}	N/A	N/A
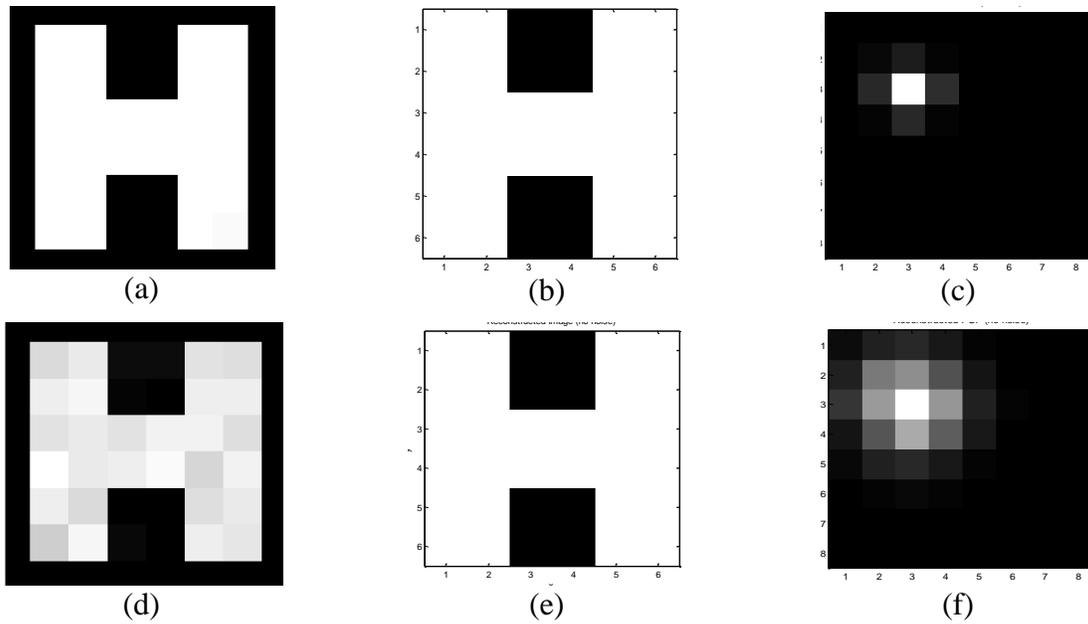
Fig. 7.37 Reconstructed weak turbulence images using (a) NAS-RIF and (b) AFBP, and (c) reconstructed PSF of weak turbulence using AFBP, and reconstructed strong turbulence images using (d) NAS-RIF and (e) AFBP, and (f) reconstructed PSF of strong turbulence using AFBP

7.6   *Adaptive Optics*

In Section 7.6, it is assumed that a simple AO system that only corrects for tip-tilt is used and the major burden of image recovery is left to the post-processing or BD module. In this section, we investigate the performance of an imaging system that exploits an AO system of variable complexity. In order to simulate the AO system, different orders of Zernike polynomials are removed from the phase screen, using a Least Square matching approach. In other words, the coefficients corresponding to tilt (Z_2 and Z_3), defocus (Z_4), coma (Z_5, Z_6), astigmatism (Z_7, Z_8), and other higher order Zernike modes in the phase screen are recovered

using a least-square sense best fitting approach. Then, the estimated phase screen is subtracted from the simulated one to resemble AO correction.

Consequently, the system PSF shrinks more and more as the number of removed Zernike modes increases. Fig. 7.38(a), (b), (c), (d), (e), (f) show the photo-detected images for a system without AO, with AO correction for Z_2 and Z_3, with AO corrections up to Z_6, with AO corrections up to Z_{10}, with AO corrections up to Z_{15}, and with AO corrections up to Z_{500}, respectively. Figs. 6 (a), (b), (c), (d), (e), (f) show the corresponding PSF values.
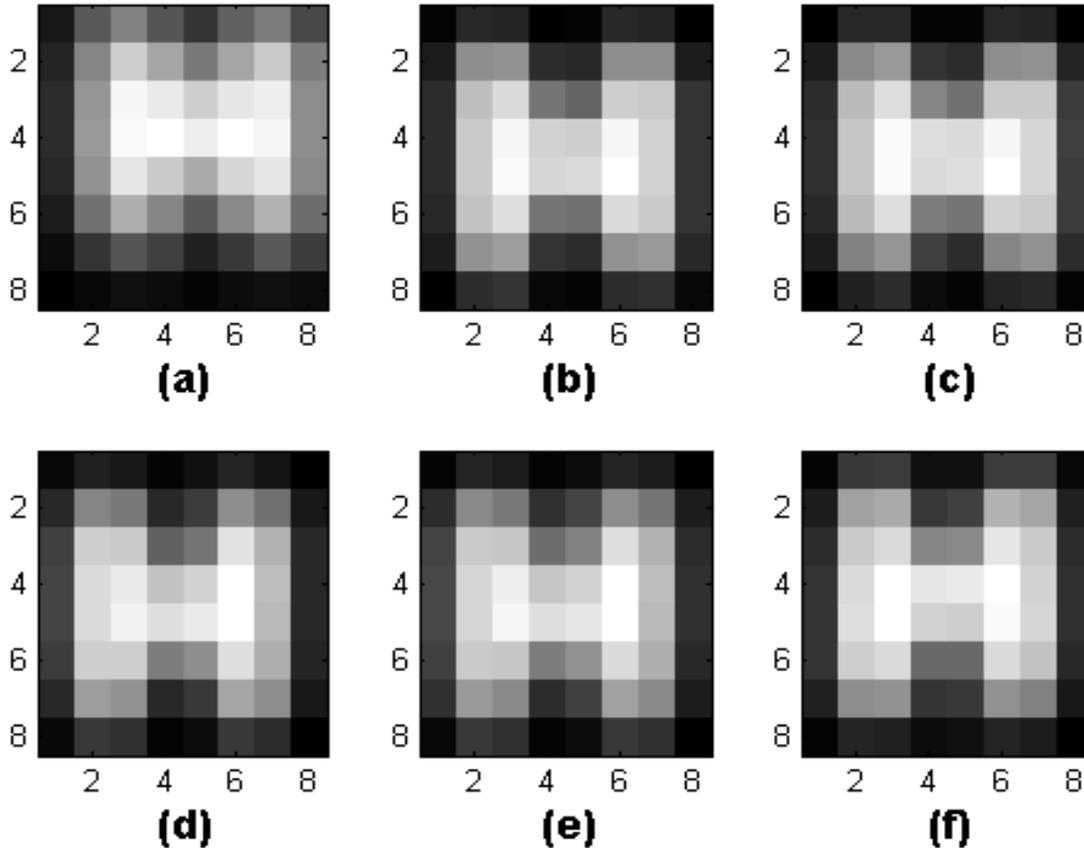
Fig. 7.38 Photo-detected images for a system (a) without AO, (b) with AO correction for Z_2 and Z_3, (c) with AO corrections upto Z_6, (d)with AO corrections upto Z_{10}, (e) with AO corrections upto Z_{15}, (f) and with AO corrections upto Z_{500}.

From Fig. 7.38 and Fig. 7.39, one can see that most performance improvements are obtained by removing the second and third Zernike modes, that is, tilt components. This was expected as tilt is responsible for 86% of the total phase error in the piston-removed phase [6]. An AO system that removes 500 Zernike modes is very complex and close to ideal. Such an AO system, almost shrinks the PSF to a delta function, i.e. diffraction limited PSF.

Table 7.8 lists SNRI, number of iterations, and MSE percentage of NAS-RIF algorithm, when AO systems of different complexities are used. The overall trend is that the number of iterations and MSE percentage decrease with the number of removed Zernike modes. At the same time, SNRI increases. However, there are some irregularities in this table. In other words, system performance is not as expected after removing 10 and 15 Zernike modes. This

might be due to incomplete implementation of the AO system, as interpolation is used to transfer from polar coordinates to Cartesian and vice versa. Moreover, since photo-detection process inherently samples the received intensity field, some inaccuracy is introduced into the system. Another reason is the fact that Zernike modes, other than tilt, have small contributions to the total phase distortion. As a result, removing lower order Zernike polynomials does not have a considerable impact on overall performance improvement. In other words, if a larger number of Zernike modes are removed from the received phase front at each stage, effect of AO corrections is more considerable and can not be suppressed by numerical inaccuracies.
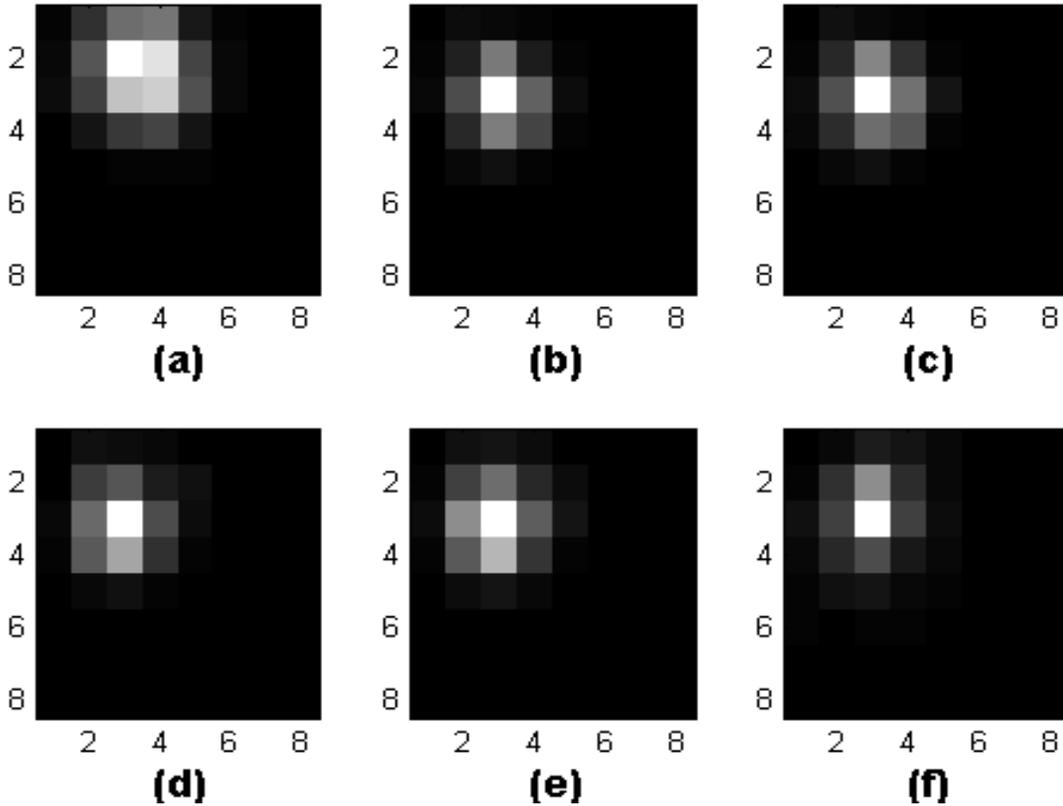
Fig. 7.39 PSF for a system (a) without AO, (b) with AO correction for Z_2 and Z_3, (c) with AO corrections upto Z_6, (d) with AO corrections up to Z_{10}, (e) with AO corrections upto Z_{15}, (f) and with AO corrections upto Z_{500}.

Table 7.8 BD performance improvement using AO.

	No AO	AO correction up to Z_3.	AO correction up to Z_6.	AO correction up to Z_{10}.	AO correction up to Z_{15}.	AO correction up to Z_{300}!
MSE% before BD	33.5%	25.3%	21.2%	19.2%	19.7%	1.7%
SNRI	2 dB	9 dB	9 dB	19.9 dB	13.7 dB	45 dB
Iteration #	165	237	174	2000	1167	576
MSE% After BD	21.2%	3.1%	2.6%	0.2%	0.8%	5×10^{-5}%

To investigate this fact, Fig. 7-40 (a), (b), (c), (d), (e), (f) show the photo-detected images for a system without AO, with AO correction for Z_2 and Z_3, with AO corrections up to Z_{55}, with AO corrections up to Z_{78}, with AO corrections up to Z_{120}, and with AO corrections up to Z_{300}!, respectively for the same received wave-front as those of Fig. 7.41 and Fig. 7.41. Moreover, Fig. 7.42 (a), (b), (c), (d), (e), (f) show the corresponding PSFs.
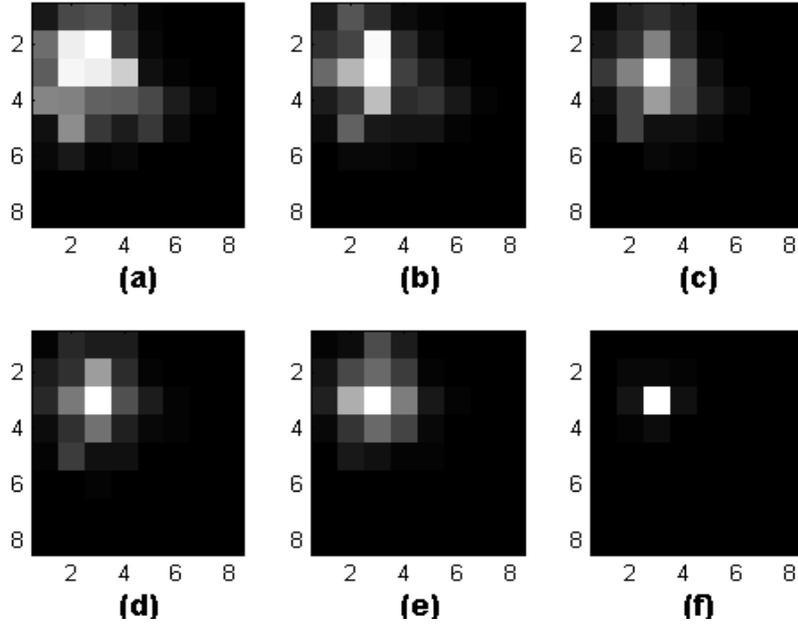
Fig. 7.40: PSF for a system (a) without AO, (b) with AO correction for Z_2 and Z_3, (c) with AO corrections upto Z_6, (d) with AO corrections upto Z_{10}, (e) with AO corrections upto Z_{15}, (f) and with AO corrections upto Z_{300}!.

Table 7.9 lists MSE percentage after AO corrections, but before BD, SNRI, number of iterations, and MSE percentage of NAS-RIF algorithm, when AO systems of different complexities are used. A monotonous trend is observed in this Table. In other words, due to removal of a large number of modes at each stage, performance improvements are not covered up by simulation inaccuracies. The only point observed in this table is the fact that SNRI of BD module after removal of 300 Zernike modes, is less than what can be obtained after removing 120 modes. As mentioned earlier, SNRI depends on the difference between MSE before and after post-processing. An AO system, capable of removing 300 Zernike modes, has already reduced the MSE percentage to 1.7 %, as a result, there is not much that can be done in the post processing stage, as image quality is very close to ideal.
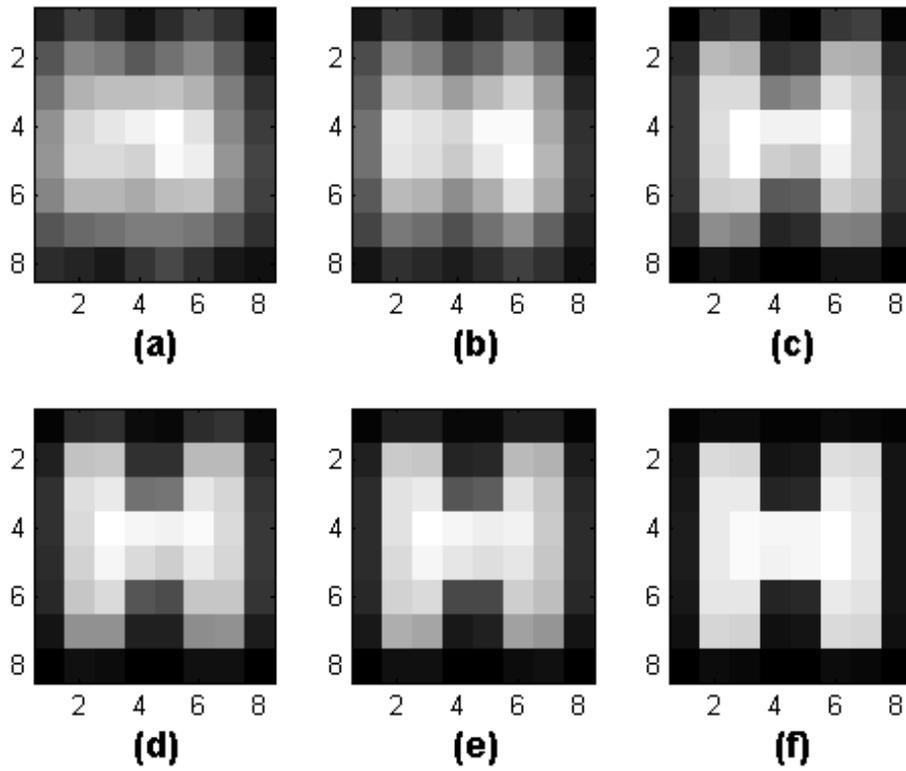
Fig. 7.41: Photo-detected images for a system (a) without AO, (b) with AO correction for Z_2 and Z_3, (c) with AO corrections upto Z_{55}, (d)with AO corrections upto Z_{78}, (e) with AO corrections upto Z_{120}, (f) and with AO corrections upto Z_{300}!.
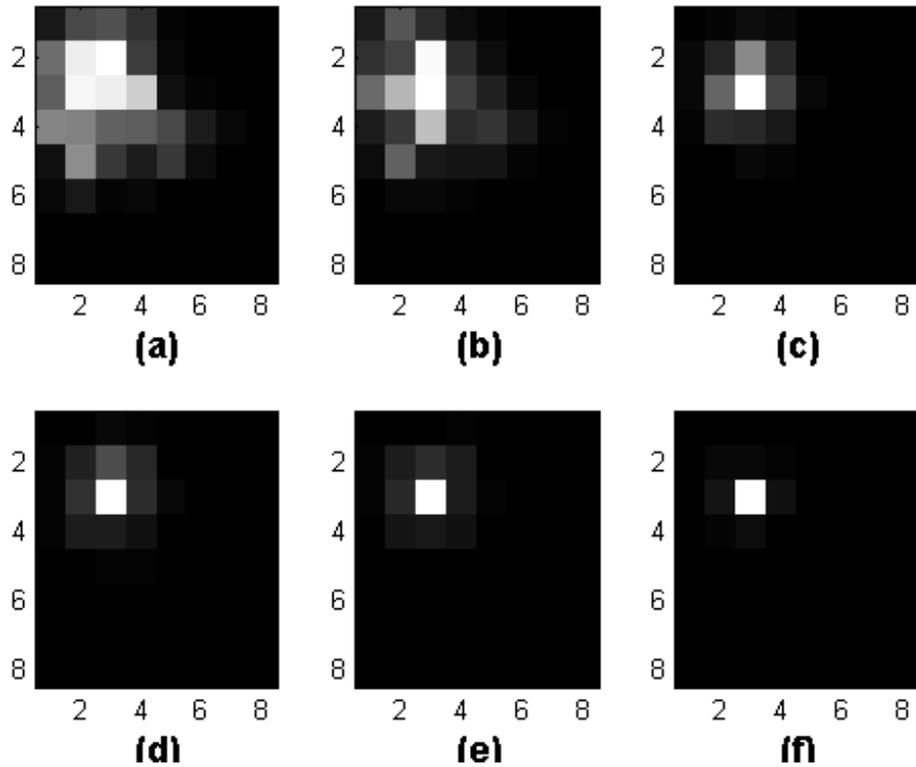
Fig. 7.42: PSF for a system (a) without AO, (b) with AO correction for Z_2 and Z_3, (c) with AO corrections upto Z_{55}, (d)with AO corrections upto Z_{78}, (e) with AO corrections upto Z_{120}, (f) and with AO corrections upto Z_{300} !.

Table 7.9: BD performance improvement with AO correction, using order Zernikes.

	No AO	AO correction up to Z_3.	AO correction up to Z_{55}.	AO correction up to Z_{78}.	AO correction up to Z_{120}.	AO correction up to Z_{300} !
MSE% before BD	33.5 %	25.3%	14%	10.8%	8.1%	1.7%
SNRI	2 dB	9 dB	29.57 dB	44.8 dB	51 dB	45 dB
Iteration #	165	237	2000	1247	1173	576
MSE% After BD	21.2%	3.14%	0.015%	3.5×10^{-4}	6.5×10^{-5}	5×10^{-5}

7.7   *Conclusion*

In Chapter 6, a spatially-multiplexed MIMO imaging system is introduced, which is inspired by MIMO multi-spot diffuse communications, first addressed by Yun and Kavehrad [7]. Though having great potential under clear weather conditions, this optical imaging system faces major limitation due to wave propagation through atmosphere, as effects of multi-scattering and turbulence cause sever degradation in the received image quality. To understand the consequences of atmospheric propagation, various simulation tools are developed at Pennsylvania State University, CICTR lab. Using these tools, an accurate

channel model is achieved; which can be used to investigate the performance of different image recovery tools.

In this chapter, several image restoration modules are explored and image quality improvements obtained using these tools are investigated assuming a simulated channel. While time-gating is used to eliminate the steady background noise caused by back-scattering, a combination of Adaptive Optics corrections and Blind Deconvolution algorithms are used to enhance the image sharpness and resolution. All these tools are proved to be very effective based on quantitative measures of image quality, such as MSE percentage and SNRI. In conclusion, the proposed spatially-multiplexed MIMO imaging system, being well equipped with image restoration modules, can provide high quality images at a faster area search rate and higher resolution compared to RF counterparts.

Chapter 8
Enabling Technologies

8.1   Introduction

Multi-Spot Diffuse Configuration (MSDC) for wireless optical communications, utilizing multi-beam transmitter and angle diversity detection, is one of the most promising ways of achieving very high capacities for use in high-bandwidth islands. Typically, the optical front-end of the receiver consists of an optical concentrator to increase the received optical signal power, and an optical bandpass filter to reject the ambient light. Using the unique properties of holographic optical elements (HOE), a novel design is proposed for the receiver optical subsystem used in MSDC. With a holographic curved mirror as an optical front-end, the receiver would achieve several dBs of improvement in the electrical signal-to-noise ratio compared to a bare photo-detector. Features as multi-functionality of the HOE, its small size, light weight, and low-cost make it a promising candidate for portable equipment in broadband free space optical links.

A parabolic holographic mirror has been recorded in silver halide at visible wavelength and its replay wavelength has been shifted to the near infrared. Employment of swelling technology resulted in permanent replay-wavelength shift without the need of hologram sealing. Despite the relatively low diffraction efficiency of holograms recorded in silver halide in principle, an improvement in the receiver signal-to-shot noise ratio of more than 20 dB has been measured. The results of the conducted experiments proved undoubtedly the great potentials of curved holographic mirrors as a key element of the receiver optical front-end in free space optical links.

8.2   Transmit Optics

For our application, we would like to have a transmitter that produces multiple beams with prescribed intensities covering the target image, which may have an asymmetric form.

There are many ways to produce multiple beams. The most straightforward approach is to have several light sources aiming at different directions. The beams can have the desired intensities and an area of any shape can be covered. However, practically, we can not have a large number of beams.

As we already mentioned, imaging photolithographic optical elements have been used at the transmitter to generate multiple beams from a single laser diode [7]. In holograms fabricated by conventional optical means, utilizing a multiple-exposure technique, an exact prescribed ratio among the intensity levels of beams can not be achieved [65]. This leads to a non-homogeneous distribution of optical power over a target image. Furthermore, this technique can not be used for very large spot arrays and for asymmetrical spot arrays. Alternatively, holograms generated by means of a computer can produce wave-fronts with any prescribed amplitude and phase distribution. Computer generated holograms (CGH) [54], see Fig. 8.1, have many useful properties. An ideal wave-front can be computed on the basis of diffraction theory and encoded into a tangible hologram. A multilevel phase CGH can have diffraction efficiency close to 100%. Holographic optical elements have insignificant physical weight and are low-cost when mass produced, *specially suited for on-board of aircrafts and spacecrafts.*
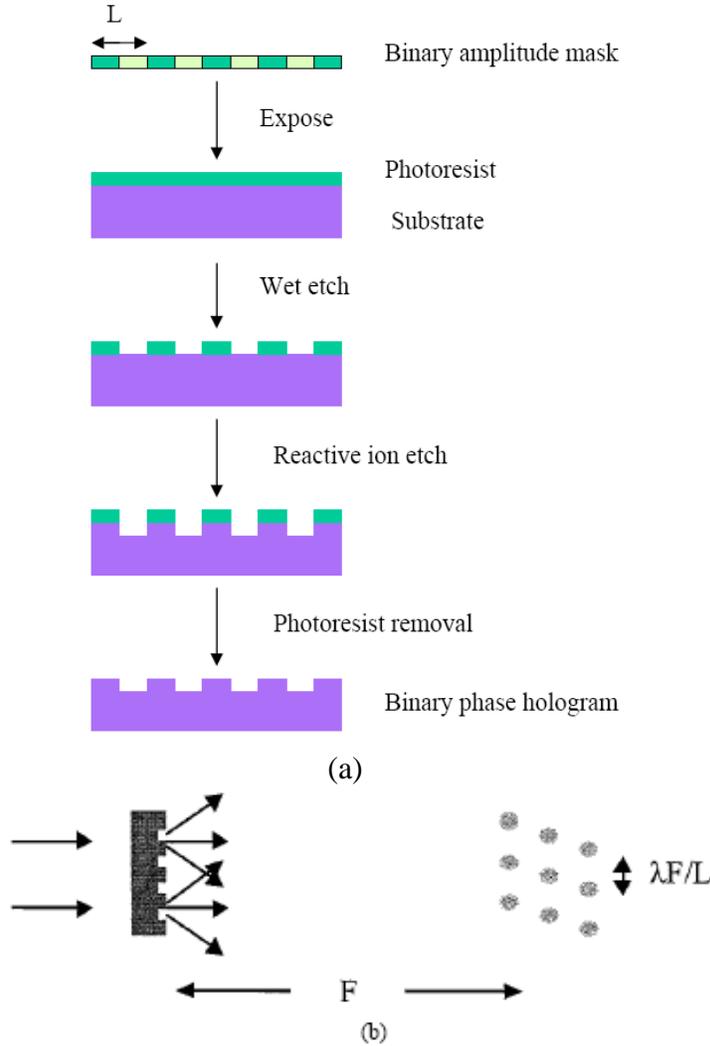
Fig. 8.1: CGH for generation of intensity-weighted spot array. (a) Fabrication. (b) Optical arrangement for far-field pattern observation.

For our application, the holographic optical element should be capable of producing multiple beams with prescribed intensities thus distributing the optical power as uniformly as possible in a given area. If the hologram consists of a great number of periodic replications of a single elementary cell (see Fig. 8.2) and is illuminated by a collimated laser beam, the far field diffraction pattern is a lattice of spots, the lattice spacing ΔS being determined by the size of the cell L [66][67] (see Fig. 8.1(b)): $\Delta S = \lambda F/L$, where λ is the wavelength of the illumination, and F is the distance between the hologram and the observation plane. Since the area that has to be covered by the spot array is very large, the size of the elementary hologram cell has to be extremely small. The amplitudes and phases of every spot are determined by the elementary cell pattern and are given by its Fourier transform modes. We are not interested in the relative phases of the generated spots. This provides more freedom in the design process and allows higher diffraction efficiency levels. To design the required photolithographic beam-splitter, the hologram elementary cell is broken up into a square array of pixels, each of

135

them imparting a specified phase delay to the incident wave-front. We employed the iterative encoding design method [66] in which the hologram cell is built up through gradual selections of changes, pixel-by-pixel, from a random initial cell pattern. The technique of simulated annealing [68] was used to minimize the cost function, defined from the difference between the desired spot pattern and the actual output pattern.

Elementary cell pattern

4x4 repeats of the elementary cell pattern

(a)

First mask *Second mask* *Third mask*

Elementary cell pattern

(b)
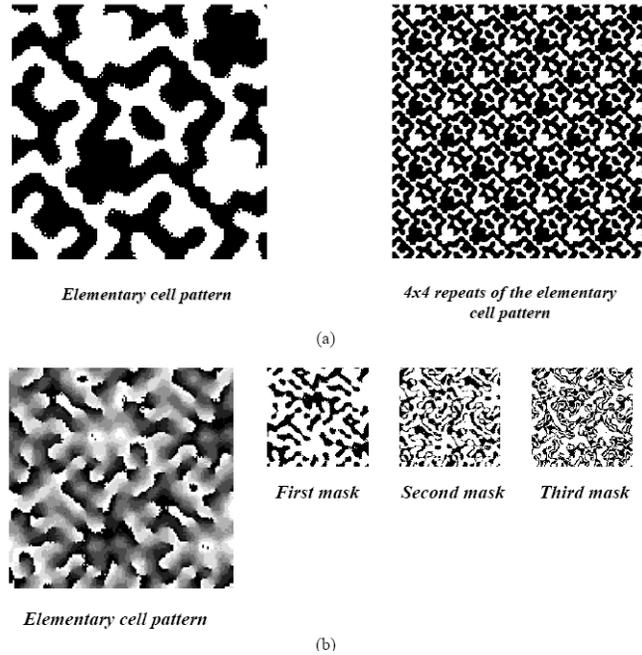
Fig. 8.2: (a) Bi-level CGH producing 8x8 beams with 80% diffraction efficiency, spot intensity variation $< 3.5\%$. (b) Eight level CGH producing 10x10 beams with 87% diffraction efficiency, spot intensity variation $< 1.55\%$.

In Fig. 8.2(a), a bi-level hologram pattern is shown. The hologram is designed to produce an array of 8x8 spots with 80% diffraction efficiency. The intensity variation between the spots is less than 3.5%. An example for an 8-level CGH producing 10x10 beams with 87% diffraction efficiency and beam intensity variation less than 1.5% is shown in Fig. 8.2(b).

The number N of the phase levels of the hologram determines the number n of the masks that have to be made: $N = 2^n$. The more phase levels, the more complexity in the fabrication of the hologram. The fabrication process of a bi-level hologram is shown in Fig. 8.1(a). From the binary computer generated amplitude mask, a surface relief binary phase hologram is produced through etching in a substrate. In the case of a multilevel hologram, this process is repeated for each of the masks. Each etch step is produced with half the depth of the previous etch step. Thus, the combination of the three etchings with the three binary amplitude masks, shown in Fig. 8.2(b), generates 8 phase levels in the final hologram. Though the bi-level hologram is easier for fabrication, it has a lower diffraction efficiency compared to the multi-level holograms, restricts the shape of the spot array to a symmetrical one, and is not capable of producing spots with prescribed but unequal intensities. If the area that has to be illuminated is, for example, L-shaped, a multilevel hologram has to be used.

136

The choice of the number of spots depends on the detection method. If wide field-of-view (FoV) receivers are to be used, the uniform illumination of the image would be best. This can be achieved with the use of a photolithographic beam-splitter that produces as many collimated beams with equal intensity as possible, thus producing a uniformly illuminated area. Such illumination will distribute the optical power uniformly over the image and will prevent the first order reflections from the surrounding reflecting objects.

If angle diversity detection is used, joint optimization of receiver FoV and the number of diffusing spots will be imperative. In principle, the more the diffusing spots are, the more optical power can be transmitted in compliance with the eye safety regulations. In addition, a more uniform distribution and a higher predictability of the channel parameters can be achieved. Since the beams emerging from the hologram are almost collimated, the eye safety limits for a point source should be observed for each one of the beams. Increasing the number of the diffusing spots up to 100 will allow increasing the intensity of each spot and of the total transmitted optical power. For larger numbers of spots, the total permissible power will remain the same, while the intensity of a single spot will decrease.

Fig. 6.15 shows a sample photolithographic 4x4 beam splitter that we have produced. The functionality was tested and demonstrated to Dr. Larry Stotts, DARPA's Deputy Director of Strategic Technology Office. The sample produced on a plastic is almost weightless.

For our application, the holographic optical element should be capable of producing multiple beams with prescribed intensities thus distributing the optical power as uniformly as possible in a given area.

Regardless of what type of detection is used, joint optimization of receiver FoV and the number of diffusing spots will be imperative. In principle, the more the diffusing spots are, the more optical power can be transmitted in compliance with the eye safety regulations. In addition, a more uniform distribution and a higher predictability of the channel parameters can be achieved. The beams emerging from the hologram are almost collimated. Increasing the number of the diffusing spots will allow increasing the intensity of each spot and of the total transmitted optical power.

For a LIDAR link employing a detector-array detection, we consider a composite receiver consisting of "N^2" branches oriented at different directions, each of them having its own photo-detector with a photosensitive area of X cm^2. All the branches are assumed to have the same field-of-view (FOV). The central branch is aimed directly towards the target image. The angle between the orientation directions of the central branch and each of the other branches is twice the field-of-view of a single branch.

As we mentioned earlier, a joint optimization of the branch field-of-view and the number of the diffusing spots on the image is necessary which has been already discussed in Section 6.3. Measured relative intensities of the hologram beam-splitter are shown in Figure 11. The measured diffraction efficiency of the sample relative to the incident light exceeded 60%.
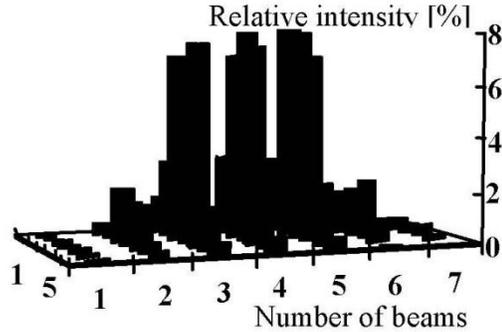
Fig. 8.3: Relative beam intensities of the multi-spot diffuser.

Beams emerging from the hologram are almost collimated. Increasing the number of the diffusing spots will allow increasing the intensity of each spot and of the total transmitted optical power. For a LIDAR link employing detector-array detection, we consider a composite receiver consisting of "N^2" branches oriented at different directions, each of them having its own photo-detector with a photosensitive area of X cm^2. All the branches are assumed to have the same field-of-view (FOV). The central branch is aimed directly towards the target image. The angle between the orientation directions of the central branch and each of the other branches is twice the field-of-view of a single branch.

8.3 Receive Optics Design

Typically, the optical front-end of the receiver consists of an optical concentrator to increase the received optical signal power, and an optical bandpass filter to reject the ambient light. In our approach, these two elements are replaced by a single holographic optical element of reflection type, called a holographic mirror. To simulate a holographic receiver optical front-end, a software package has been developed using ray tracing techniques. Three different types of holographic mirrors have been considered: spherical mirror on a spherical substrate, spherical mirror on a flat substrate, and parabolic mirror on a flat substrate.

We have investigated the characteristics and performance of optical interference filters. Typically, interference filters with a narrowband spectral response have a low transmittance and reduce the signal by 3 dBo. These filters work well with collimated radiation which is not an option in our application. The narrowband interference filters are sensitive to incident angle of radiation and the peak wavelength shifts toward shorter wavelengths with increasing the incident angle. The filter can be designed with a shift toward longer wavelengths in order to accommodate the spectral shift by an incident angle. Given a diffuse radiation pattern, however, performance further deteriorates because of non-collimated radiation, i.e., the incident angle varies from normal incidence to the maximum angle determined by the field-of-view (FOV) of the receiver lens. When the spectral response of the filter is averaged over all incident angles, the full-width-half-maximum (FWHM) is broadened and the signal is further attenuated. The maximum transmittance is further reduced for non-polarized light; the case in imaging through cloud/fog configuration. Since the laser diode and the filter wavelengths are specified by the manufacturer within given tolerances, and there is also a wavelength drift due to temperature changes, a narrow-band interference filter does not seem to be practical. In our measurements, we used an interference filter with a

FWHM equal to $\delta\lambda = 70$ nm with an internal transmittance T = 0.60 that can accommodate the shift to shorter wavelengths and has insignificant changes in its performance over a wide range of incident angles. We also used a colored glass filter that is independent of the incident angle and of the polarization with a $\delta\lambda = 300$ nm and a high internal transmittance of T= 0.95.

Practically, the FOV of the receiver lens is determined by those marginal rays incident at an angle θ_{max} that pass through the lens and hit the edge of the detector, where θ_{max} depends on the refractive index of the lens material and on the ratio of the "lens and the detector radii". The FOV is restricted by: (i) total internal reflection that occurs on boundaries from a medium of a higher refractive index to a medium with a lower index and (ii) by a small ratio of the lens and detector radii. We need to avoid a large detector area to avoid background noise and excessively-delayed photons. As a design example, assume the detectors used in the link have a radius $r_{det} = 2$ mm or 5 mm. At the same time, as predicted by our theoretical evaluations and confirmed by the measurements, the performance of optical filter deteriorates when a large cone of rays pass through it. Therefore, a lens with a smaller FOV works better in combination with an interference filter. With a plano-convex lens of a focal length f = 24 mm and a diameter d = 24 mm and $\frac{r_{lens}}{r_{det}} = 2.4$ with the colored glass filter and a FOV≈50 degrees, we measured an improvement in the path loss of about 8 dBo over a distance of 2 m, compared to the case of a stand-alone detector. When used indoors, the fluorescent light has very little effect on the SNR and the desk-top light deteriorates it by no more than 2 dBo. We also used an aspheric lens of a focal length f = 79mm and d = 80mm. This lens, although bulky, has a FOV ≈ 35 degrees and works well with the interference filter. We measured an overall improvement in the detected power of 3 dBo. With this combination, the fluorescent and the desk-top lights have no effect even as close as 30 cm away from and 10 cm above the detector. The aspheric lenses are suitable for short focal lengths and reduced spherical aberrations and allow more energy to be concentrated on the small detector area. These are advantageous for high power-throughput applications. We used index-matching liquid between the optical components in the receiver to reduce Fresnel reflection losses due to refractive-index mismatch. According to our theoretical evaluations, up to 2.9 dBo of the optical power reaching the detector can be lost because of Fresnel reflection. Therefore, anti-reflection coatings on all components including the detector become a necessity for an optimized system.

As stated earlier, the light reaching the entrance aperture of the optical front-end is filtered in order to reduce the optical noise and is concentrated in order to increase the irradiance onto the photodiode. Using the unique properties of holographic optical elements (HOE), a novel design is proposed for the receiver optical subsystem used in the proposed imaging system. With a holographic curved mirror as an optical front-end, the receiver would achieve several dBs of improvement in the electrical signal-to-noise ratio compared to a bare photo-detector. The main advantages of HOEs over conventional systems that consist of a lens concentrator and an optical filter are multi-functionality, independence of their physical configuration, insignificant weight, low cost, etc.

Hence, these two elements, i.e., the filter and concentrator are replaced by a single holographic optical element of reflection type, called a holographic mirror.

8.3.1 Holographic Mirrors

The holographic optical elements have several unique properties. They are spectrally selective. Their optical configuration is independent of their substrate configuration. For example, a flat reflection holographic optical element can be made to function as a plane, spherical, hiperboloidal, paraboloidal or ellipsoidal mirror. A single hologram can perform multiple functions. The diffraction efficiency of a volume phase hologram can be close to 100%. The physical weight of a holographic optical element is insignificant compared to conventional optics. These are low-cost in mass production.

The higher spectral selectivity and the lower angular selectivity of reflection holograms compared to transmission holograms make these more suitable for implementing the functions needed for receiver optical front-end.

Volume reflection holograms are called holographic mirrors, though they do not reflect light as conventional mirrors. They diffract light obeying Bragg law, which establishes a relation between the spacing of the planes of diffraction Λ, the wavelength of light λ and the half-angle $\theta \ \square$ between the incident and the diffracted beams:

$$2\,n\,\Lambda\,\sin\theta = \lambda \tag{8.1}$$

where n is the refractive index of the material.

If θ_1 and θ_2 are values of the angles of incidence and diffraction which satisfy Bragg condition for a wavelength λ, the diffraction efficiency of a lossless reflection hologram is [69]:

$$\eta = \left(1 + \frac{1 - \xi^2/\mu^2}{\sinh^2\sqrt{\mu^2 - \xi^2}}\right)^{-1} \tag{8.2}$$

where $\mu = \dfrac{k\,d}{\sqrt{|c_R c_S|}}$, $\quad k = \dfrac{\pi\,\Delta n}{\lambda}$, $\quad c_R \approx \cos\theta_1$, $\quad c_S \approx \cos\theta_2$.

Δn is the amplitude of the refractive index grating, d is the thickness of the hologram, $\xi = \dfrac{d\,\Omega}{2\,c_S}$ is the off-Bragg parameter that accounts for the deviations from the Bragg condition,

$$\Omega = \beta\,\Delta\theta\,\sin(\theta_1 - \theta_2) + \Delta\beta\left[1 - \cos(\theta_1 - \theta_2)\right] \tag{8.3}$$

$\beta = \dfrac{2\,\pi\,n}{\lambda}$ is the magnitude of the wave vectors satisfying the Bragg condition and $\Delta\beta$ is the deviation from that value due to a change in the reconstruction beam wavelength, and $\Delta\theta$ is the deviation from the Bragg angle θ at the reconstruction.

It can be seen from (8.3) that changes in the angle of incidence and wavelength from the recording conditions may be mutually compensatory and at the reconstruction, Bragg condition may be satisfied for angles and wavelengths that differ from those at the recording. This is illustrated in Fig.-4, where the diffraction efficiency dependence on the angle and wavelength is shown for a reflection hologram recorded at 850nm by two counter-propagating

plane waves at normal incidence, i.e., $\theta_1 = 0$ and $\theta_2 = \pi$. The influence of the amplitude Δn of the refractive index grating and the hologram thickness d on the peak diffraction efficiency, angular and spectral selectivity of the hologram is illustrated on Fig. 8.4, Fig. 8.5 and Fig. 8.6.

(a)

(b)
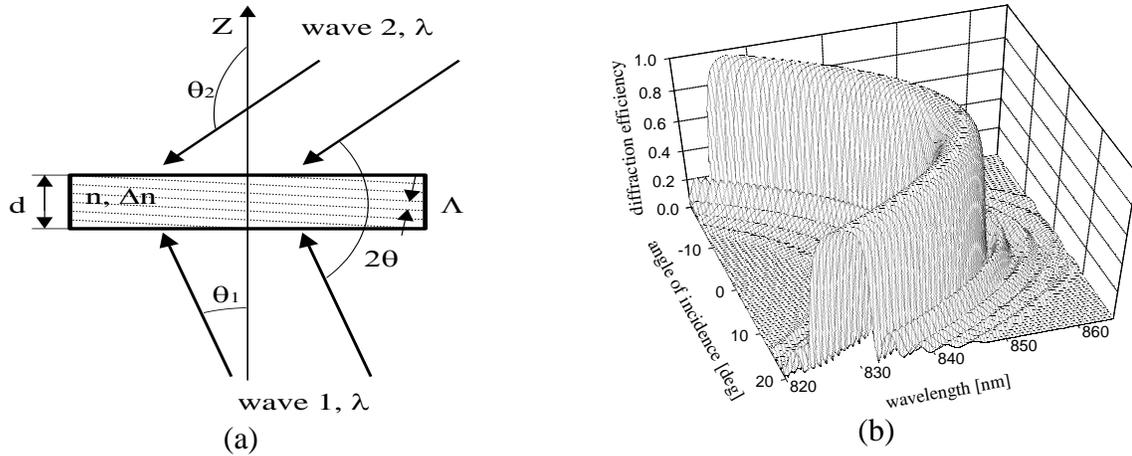
Fig. 8.4. (a) Reflection hologram recorded by two plane waves. (b) Angular-spectral selectivity for $\theta_1 = 0$, $\theta_2 = \pi$, $\lambda = 850\text{nm}$, $n = 1.5$, $\Delta n = 0.01$, $d = 60\mu\text{m}$.

(a)

(b)
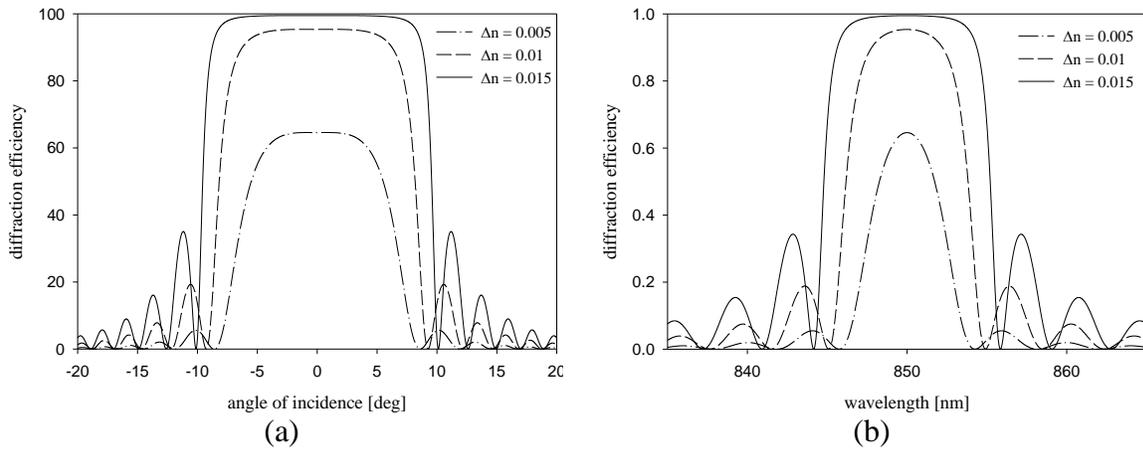
Fig. 8.5. (a) Angular and (b) spectral selectivity of a reflection hologram ($\theta_1 = 0$, $\theta_2 = \pi$, $\lambda = 850\text{nm}$, $n = 1.5$, $d = 60\mu\text{m}$) for different amplitudes of the refractive index grating.
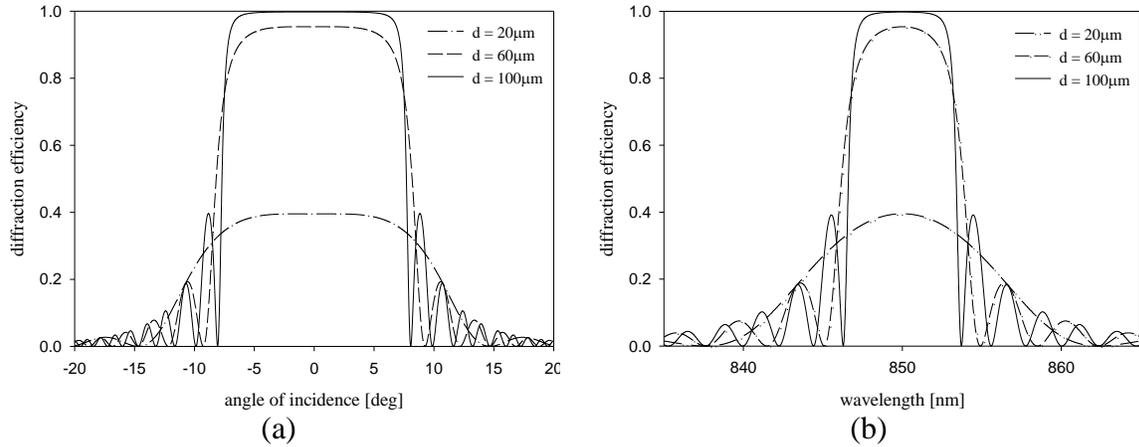
Fig. 8.6. (a) Angular and (b) spectral selectivity of a reflection hologram ($\theta_1 = 0$, $\theta_2 = \pi$, $\lambda = 850\text{nm}$, $n = 1.5$, $\Delta n = 0.01$) for different thickness of the recording medium.

If at recording, one or both plane waves are replaced by a spherical wave, the recorded reflection hologram will exhibit concentrating features like conventional curved mirrors regardless of the geometry of the substrate. In this section, we consider a spherical holographic mirror recorded by two spherical waves with $\lambda = 850\text{nm}$ on a spherical substrate in Fig. 8.7. The recording medium is assumed to be 100 μm ☐ thick and to have a refractive index of 1.5 (for example, dichromated gelatin). The amplitude of the refractive index grating is assumed to be $\Delta n = 0.025$.
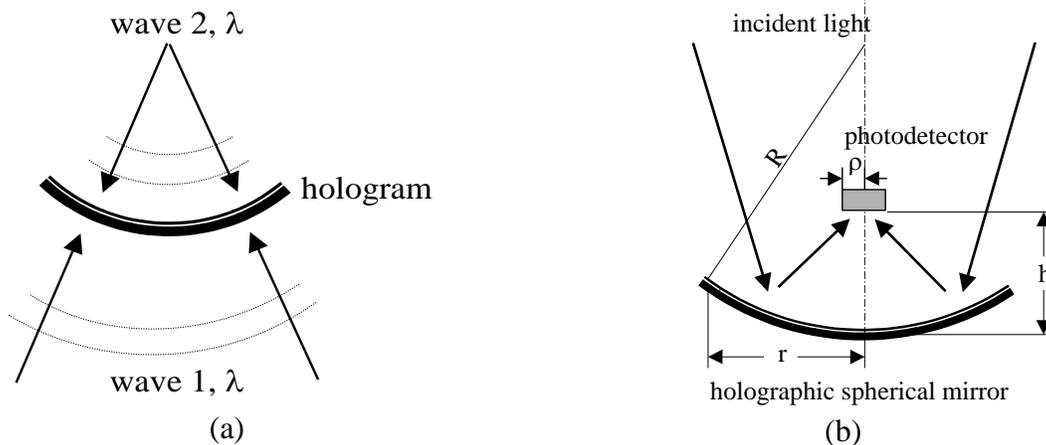
Fig. 8.7. (a) Recording of a spherical holographic mirror on a spherical substrate. (b) A holographic spherical mirror as a receiver optical front-end.

8.3.2   Holographic Spherical Mirror as a Receiver Optical Front-End

Angle diversity detection is based on signal reception by several narrow field-of-view receiver components. Each component could be a separate branch consisting of photo-detector with its own optics [72], or a pixel from an array of photo-detectors sharing common optics [7][73][74]. Holographic optical elements can be employed with both approaches. In

this report, we investigate a holographic curved mirror to be used on the branches of a composite receiver. Our previous investigation [52] showed promising results for the system performance when the receiver consists of 7 branches, each of them with a $\text{FOV}=11.5°$. Since in [52] bare photo-detectors were considered, we expect further improvements in the system performance when an optical concentrator and a narrow-band filter are used.

Let us consider a holographic spherical mirror (HSM) combined with a photo-detector (Fig. 8.7 b) in a single system. The HSM will perform concentrating function as a conventional spherical mirror. Furthermore, due to its nature, it will diffract light from a very narrow spectral range toward the detector, thus performing filtering function, as well. The characteristics of the system, such as field-of-view (FOV), concentration ratio, spectral bandwidth, etc., can not be derived directly from the hologram characteristics. The performance of the system depends also on the geometric configuration, i.e., size and curvature of the HSM, size and position of the photo-detector.

8.3.3 *Field-of-view and spectral bandwidth*

Since the considered receiver optics design is intended to be used with a multi-spot diffusing configuration, we assume the optical signal is directional, not isotropic. In fact, if the diffusing spot covers an area of 25cm^2 on the image, it will subtend an angle of about 1° at a distance of 3m. So, we assume that the signal power reception depends only on the angle of incidence and not on the receiver FOV. Also, we assume ambient light is isotropic within the receiver FOV, i.e., its reception depends on the FOV and not on the angle of incidence. With these assumptions, the detected signal radiant flux is:

$$\Phi_s(\varphi,\lambda) = \int_{r=0}^{r}\int_{\psi=0}^{2\pi} E_s\, \eta(\varphi,r,\psi;\lambda)\frac{r}{\sqrt{1-(r/R)^2}}\cos\varphi\,d\psi dr = E_s\,A_{s,\text{eff}}(\varphi,\lambda) \tag{8.4}$$

where $E_s\,[\text{W/cm}^2]$ is the signal irradiance and $A_{s,\text{eff}}(\varphi,\lambda)[\text{cm}^2]$ is the signal effective area of the receiver. The diffraction efficiency η of the HSM is given by (8.2) and depends not only on the angle of incidence φ and the wavelength λ, but also on the particular point (r,ψ) of the hologram, since the deviation from the Bragg angle would be different at different points of the hologram.

The received ambient light radiant flux is:

$$\Phi_{bg} = \int_{\lambda=0}^{\infty}\int_{\phi=0}^{\pi/2}\int_{r=0}^{r}\int_{\psi=0}^{2\pi} L_{bg}\,\eta(\phi,r,\psi;\lambda)\frac{r}{\sqrt{1-(r/R)^2}}\,2\pi\sin\phi\cos\phi\,d\psi dr d\phi d\lambda$$

$$= L_{bg}\int_{\lambda=0}^{\infty}(AS)_{bg,\text{eff}}\,d\lambda \tag{5}$$

(8.5)

where $L_{bg}\,[\text{W/cm}^2/\text{sr/nm}]$ is the spectral background radiance and $(AS)_{bg,\text{eff}}(\lambda)[\text{cm}^2\text{sr}]$ is the effective area-solid angle product of the receiver which accounts

for the receiver effective area for the ambient light and the effective solid angle within which the ambient light is received.

It is clear that the receiver would have different angular response at different wavelengths due to the specific nature of the HSM (see Fig. Fig. 8.4(b)). We define the field-of-view of the receiver to be the angle at which the signal effective area, i.e., the effective area at the signal wavelength, becomes a half of its maximum value. Using ray tracing, we calculated the signal effective area at different angles of incidence for different geometric configurations of the receiver. The results are plotted in Fig. 8.8. It can be seen that the maximum value and the angular dependence of the signal effective area is strongly influenced by each of the geometrical factors, i.e., the aperture radius r of the HSM, the radius R of its curvature, the radius ρ $\Box$ of the photo-detector and the distance h between the HSM and the detector.

Similarly, the receiver would have different spectral response at different angles of incidence (Fig. 8.4(b)). Thus, in the case of non-isotropic ambient light, the receiver band strongly depends on the directionality of the ambient light sources. We define the spectral bandwidth of the receiver to be the full width at half maximum of the spectral dependence of its effective area-solid angle product $(AS)_{bg,eff}(\lambda)$ (see Eq.-5). This value determines the ambient light reception in the case of isotropic background. In Fig. 8.9, the spectral dependence of $(AS)_{bg,eff}(\lambda)$ is shown for the same geometrical configurations of the receiver as in Fig. 8.8.

In Fig. 8.10(a), the effective area of the receiver is plotted against the angle of incidence for different signal wavelengths for a fixed geometrical configuration. This system achieves a field-of-view of about $11.5°$. The bandwidth of the same system for an isotropic ambient light is about 24nm, see Fig. 8.10(b).
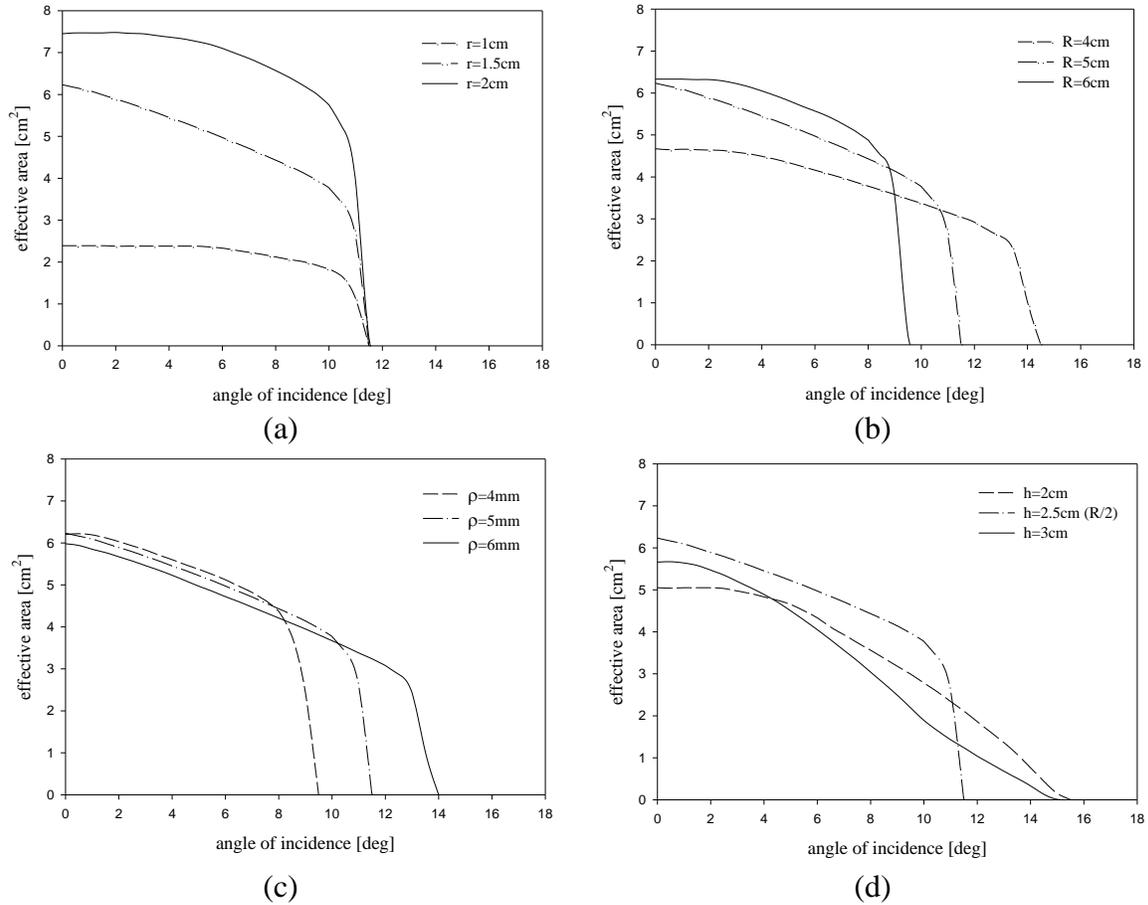
Fig. 8.8. Angular dependence of the receiver effective area at the wavelength for which $(AS)_{bg,eff}$ takes maximum value. (a) R=5cm, ρ=5mm, h=R/2; (b) r=1.5, ρ=5mm, h=R/2; (c) R=5cm, r=1.5, h=R/2; (d) R=5cm, r=R/2, ρ=5mm.
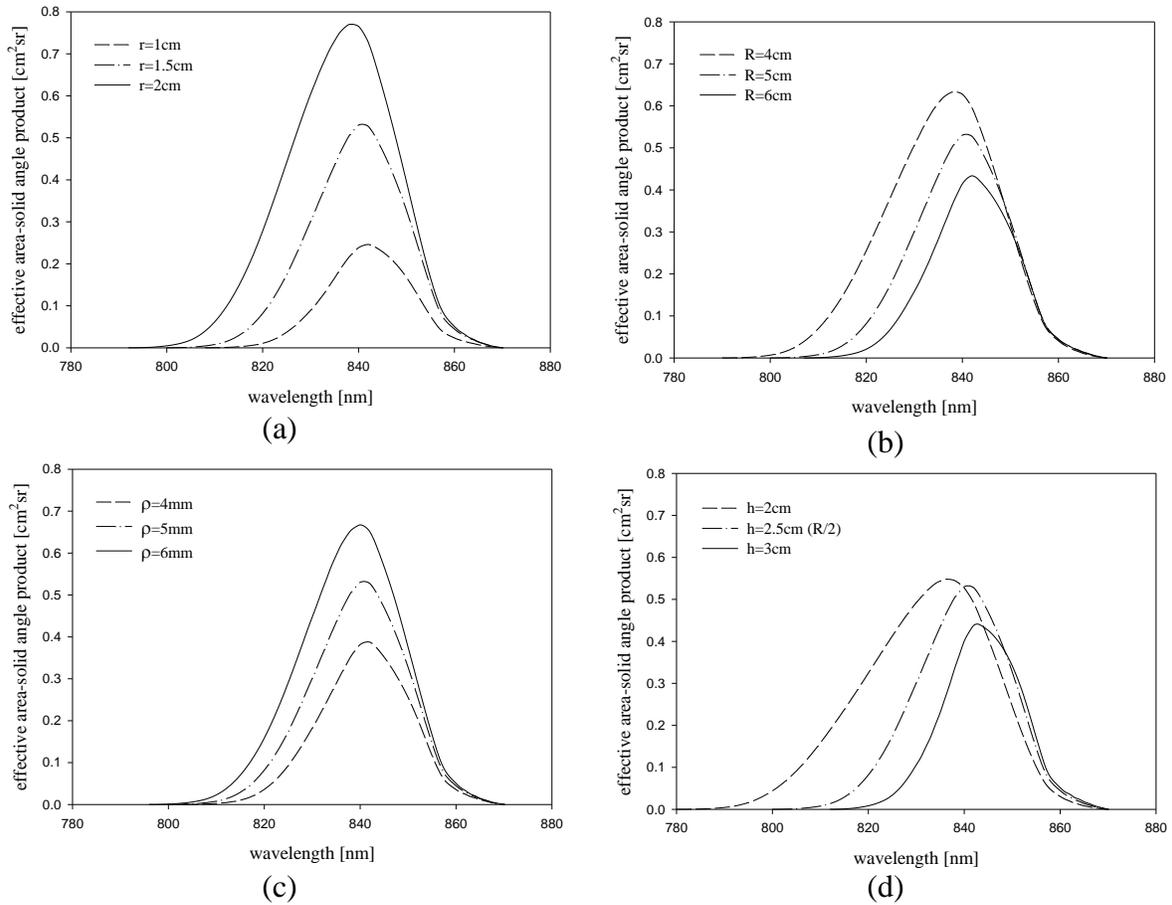
Fig. 8.9. Spectral dependence of the receiver effective area-solid angle product for different geometrical configurations. An isotropic distribution of the optical power is assumed. (a) R=5cm, ρ=5mm, h=R/2; (b) r=1.5, ρ=5mm, h=R/2; (c) R=5cm, r=1.5, h=R/2; (d) R=5cm, r=R/2, ρ=5mm.
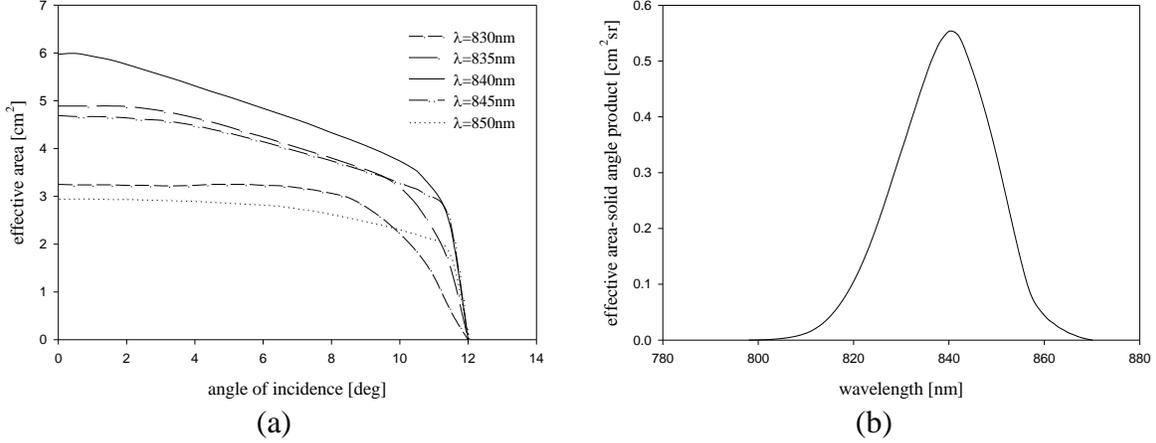
Fig. 8.10. (a) Signal effective area for different signal wavelengths and (b) spectral response of a holographic spherical mirror for an isotropic ambient light (r=1.5cm, R=4.8cm, ρ=5mm, h=2.4cm).

8.3.4 *Signal gain and figure-of-merit*

Let us consider an ideal optical concentrator (IC) combined with an interference filter (IF). An ideal concentrator would collect at its exit aperture all the rays that strike its entrance aperture within its field-of-view. Provided that the interference filter bandwidth $\Delta\lambda$ is wide enough to ensure maximum signal transmission all over the concentrator field-of-view, in the case of a directional optical signal, the received signal radiant flux would be:

$$\Phi_s(\varphi) = E_s \, T \, A_{in} \, \cos\varphi = E_s \, A_{s,eff}^{id}(\varphi) \tag{8.6}$$

where T is the peak transmittance of the interference filter and A_{in} is the area of the entrance aperture. For isotropic ambient light, the received background radiant flux would be:

$$\Phi_{bg} = L_{bg} \, T \, A_{in} \, 4\pi\sin^2\!\left(\frac{FOV}{2}\right)\Delta\lambda = L_{bg} \, (AS)_{bg,eff}^{id} \, \Delta\lambda \tag{8.7}$$

The optical signal-to-noise ratio (SNR) of the concentrator-filter system depends only on the field-of-view and the spectral bandwidth of the system:

$$SNR_{opt}^{id}(\varphi) = \frac{\Phi_s(\varphi)}{\Phi_{bg}} = \frac{E_s}{L_{bg}} \frac{\cos\varphi}{4\pi\sin^2\!\left(\dfrac{FOV}{2}\right)\Delta\lambda} \tag{8.8}$$

while the electrical SNR depends on the size of the entrance aperture and the filter transmittance, as well:

$$SNR_{el}^{id}(\varphi) \propto \frac{\Phi_s{}^2(\varphi)}{\Phi_{bg}} = \frac{E_s{}^2}{L_{bg}} \frac{T \, A_{in} \, \cos^2\varphi}{4\pi\sin^2\!\left(\dfrac{FOV}{2}\right)\Delta\lambda} \tag{8.9}$$

147

Thus, a concentrator-filter system may exhibit a worse optical SNR but a better electrical SNR compared to other similar systems. For this application, we are interested in electrical signal-to-noise ratio, so we will refer to it simply as signal-to-noise ratio and we define a figure-of-merit for the IC-IF system as

$$M^{id}(\varphi) = \frac{T\,A_{in}\cos^2\varphi}{4\pi\sin^2\left(\dfrac{FOV}{2}\right)\Delta\lambda} \tag{8.10}$$

In the case of a HSM, the signal-to-noise ratio and the figure-of-merit are:

$$SNR(\varphi) \propto \frac{\Phi_s^{\,2}(\varphi)}{\Phi_{bg}} = \frac{E_s^{\,2}}{L_{bg}}\frac{A_{s,eff}^{\,2}}{\displaystyle\int_{\lambda=0}^{\infty}(AS)_{bg,eff}\,d\lambda} \tag{8.11}$$

and

$$M(\varphi) = \frac{A_{s,eff}^{\,2}}{\displaystyle\int_{\lambda=0}^{\infty}(AS)_{bg,eff}\,d\lambda} \tag{8.12}$$

To be able to judge the quality of the receiver optical front-end, let us compare the performance of a receiver equipped with optics to the performance of a bare photo-detector. We assume the photo-detector has the same field-of-view as the concentrator, its photosensitive area is equal to the concentrator exit aperture area A_{out}, and its responsivity is constant over 200nm spectral interval centered around the signal wavelength and is zero out of this range.

We define the signal gain to be the increase of the received signal radiant flux compared to the radiant flux received by the bare photo-detector. For a receiver front-end consisting of an ideal concentrator and an interference filter, the signal gain is:

$$G_s^{id} = 10\log\left(\frac{E_s\,T\,A_{in}\cos\varphi}{E_s\,A_{out}\cos\varphi}\right) = 10\log\left(\frac{T\,A_{in}}{A_{out}}\right) \tag{8.13}$$

and for a HSM serving as receiver optics:

$$G_s(\varphi) = 10\log\left(\frac{\displaystyle\int_{r=0}^{r}\int_{\psi=0}^{2\pi}E_s\,\eta(\varphi,r,\psi;\lambda)\dfrac{r}{\sqrt{1-(r/R)^2}}\cos\varphi\,d\psi dr}{E_s\,A_{out}\cos\varphi}\right) = 10\log\left(\frac{A_{s,eff}}{A_{out}\cos\varphi}\right) \tag{8.14}$$

Similarly, we define the figure-of-merit gain as the figure-of-merit improvement compared to bare photo-detector. In the case of IC-IF it is:

$$G_M^{id} = 10 \log \left(\frac{T A_{in} \cos^2 \varphi}{4\pi \sin^2 \left(\frac{FOV}{2} \right) \Delta\lambda} \frac{4\pi \sin^2 \left(\frac{FOV}{2} \right) 200}{A_{out} \cos^2 \varphi} \right) = 10 \log \left(\frac{200\, T\, A_{in}}{\Delta\lambda\, A_{out}} \right) \qquad (8.15)$$

In the case of a HSM the figure-of-merit gain is:

$$G_M(\varphi) = 10 \log \left(\frac{A_{s,eff}^2}{\int_{\lambda=0}^{\infty} (AS)_{bg,eff}\, d\lambda} \frac{4\pi \sin^2 \left(\frac{FOV}{2} \right) 200}{A_{out} \cos^2 \varphi} \right) \qquad (8.16)$$

Fig. 8.11 shows the signal gain and the figure-of-merit gain of the same HSM-photo-detector system compared to two combinations IC-IF.
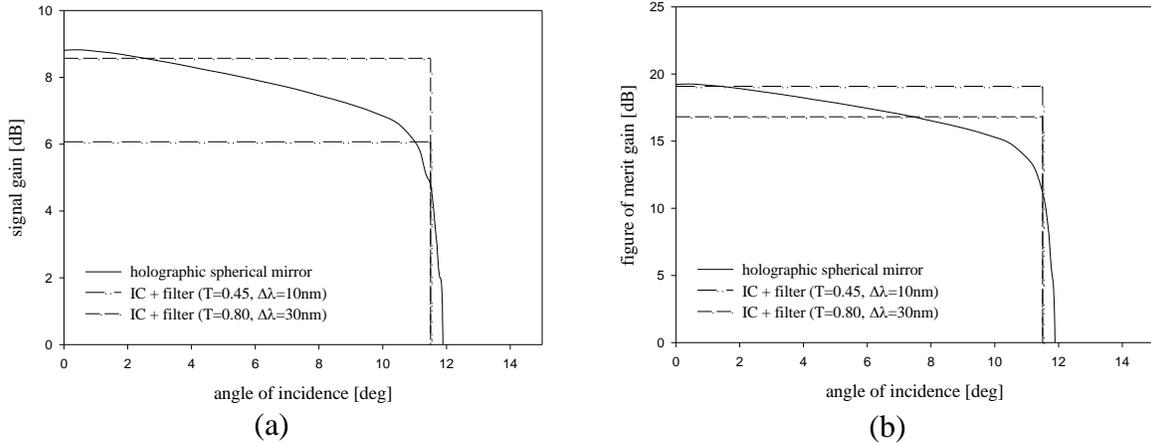
Fig. 8.11. (a) Signal gain and (b) figure-of-merit gain for a holographic spherical mirror and an ideal concentrator combined with different interference filters.

The ideal concentrator is assumed to have the same FOV and entrance aperture as the HSM, and its exit aperture is matched to the same photo-detector as in the case of a HSM. The interference filter in the first IC-IF combination has a peak transmittance T=45% and a bandwidth $\Delta\lambda$ = 10nm. A spectral bandwidth of 10nm would be enough to ensure that the optical signal is always received at the maximum transmittance of the filter for a $FOV = 11.5°$, and the typical peak transmittance of the commercially available interference filters is about 45%. The second IC-IF combination employs an interference filter with a higher transmittance (T=80%) but larger bandwidth ($\Delta\lambda$=30nm). It was chosen to show that the effect of the better signal reception (due to a higher transmittance) can be totally suppressed by the worse ambient light rejection (due to a larger bandwidth). This can be clearly seen in Fig. 8.8. Naturally, the optical system consisting of an ideal concentrator and an interference filter with a high peak transmittance achieves the highest signal gain. However, with respect to SNR, i.e., the figure-of-merit gain, the same system has the worst performance for almost over the entire field-of-view range of values.

At normal incidence, the HSM is performing better than both combinations IC-IF, achieving figure-of-merit gain of more than 19dB. The degradation of its performance near the edge of the field-of-view is in principle due to the non-ideal nature of the spherical mirror as an optical concentrator [71]. Nevertheless, HSM still provides more than 11dB improvement in the electrical signal-to-noise ratio compared to a bare photodiode.

Apparently, the receiver can achieve a better performance if the optical front-end consists of an ideal concentrator and an appropriate interference filter than if a HSM is employed. However, when designing the receiver optical front-end, one should not just consider the system performance. For portable applications, issues as the system volume, weight and cost, as well, must be taken into account. There exists real optical concentrators exhibiting nearly ideal characteristics and the compound parabolic concentrator (CPC) is one of them. A research group at Berkeley investigated the possibility of employing such concentrators for the receiver [72][75]. Unfortunately, a CPC has its own drawbacks. First, the concentrator is very long. For example, for an entrance aperture radius of 1.5cm and a $FOV = 11.5°$, the length of a hollow CPC is 8.85cm. Also, since the CPC is a conventional optical concentrator, it has to be equipped with an interference filter, increasing the total price of the receiver optics. On the contrary, the holographic optical elements have a small size and weight compared to the conventional optical elements. Also, these are low-cost when mass-produced.

As stated earlier, a parabolic holographic mirror is being recorded in silver halide at visible wavelength and its replay wavelength is being shifted to the near infrared. Swelling technology results in a permanent replay-wavelength shift without the need for hologram sealing. It is expected to achieve a significant improvement in the receiver signal-to-shot noise ratio in the order of tens of dBs. With respect to ambient light, even a weak diffused background light is much stronger than the optical signal. To improve the signal-to-noise ratio, an optical filter that would efficiently reject the optical noise is needed. Thus, functionally, receiver branch optical front-end consists of an optical concentrator, an optical band-pass filter and a photodiode (Fig. 7a). The light reaching the entrance aperture of the optical front-end is filtered in order to reduce the optical noise and is concentrated in order to increase the irradiance onto the photodiode. Physically, these two functions can be performed by a single holographic optical element (Fig. 8.12(b)). The main advantages of holographic optical elements (HOE's) over conventional systems that consist of a lens concentrator and an optical filter are multi-functionality, independence of their physical configuration, insignificant weight, low cost, etc.
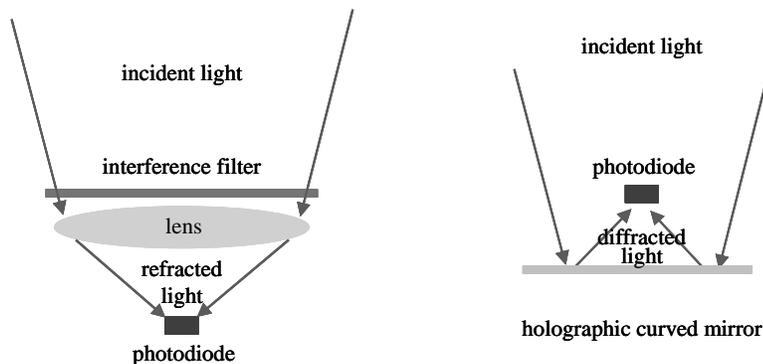
Fig. 8.12. (a) Refractive and (b) holographic receiver optical front-ends.

The higher spectral selectivity and the lower angular selectivity of reflection holograms compared to transmission holograms make these more suitable for implementing the functions needed for receiver optical front-end [69]. If at the recording, one or both recording waves are spherical, the recorded reflection hologram (holographic mirror) will exhibit concentrating features like conventional curved mirrors regardless of the geometry of the substrate. Furthermore, due to its inherent spectral selectivity, the holographic mirror will diffract light from a very narrow spectral range, thus performing filtering function, as well. In our application, holographic optical elements have additional advantage of having a light weight.

In our previous works, we have theoretically investigated in details the performance of holographic spherical mirrors (HSM) as receiver optical front-end, and we have presented a theoretical comparative study of HSM and holographic parabolic mirrors (HPM). The results of this study suggest that HPM is a better choice for small field-of-view receiving elements [76].

8.3.5 *Holographic Parabolic Mirror Fabrication*

There are no materials sensitive to the near infrared and suitable for volume reflection holography. Since the holographic optical element is to be used at near infrared wavelengths, the hologram has to be recorded at visible wavelength and then the medium properties have to be altered in order to shift the replay wavelength to the near infrared.

Two holographic mirrors with a 30mm diameter were recorded. The first one is an on-axis parabolic mirror with a focal length of 22 mm and the second one is an off-axis parabolic mirror with the focal point positioned 65 mm from the hologram center on an axis tilted at about 60 degrees with respect to the hologram surface. The second holographic optical element has the advantage that its focal point is outside its entrance aperture and a photodiode positioned there would not shadow the hologram. Fig. 8.13 (a) and (b) depict the experimental arrangement used for holographic mirrors recording.

The two holograms were recorded in silver halide film HP-650, production of CLOSPI [77], with a He-Ne laser at 632.8 nm. The holograms are of a mixed amplitude/phase type, thus having relatively low diffraction efficiency. After recording, the holographic emulsion has to be swelled in order to shift the replay wavelength to the near infrared. The researchers at CLOSPI have developed such technology in the 80-ties [78][79]. Recently, however, this technology has been further developed so that the swelled holograms are practically insensitive to humidity and there is no need for hologram sealing. Only a weak shrinkage of the swelled hologram follows in the first few weeks after the swelling. No further change of the holographic emulsion properties occurs afterward. The off-axis HPM combined with a photodiode with a 6 mm diameter provided a receiver acceptance angle of 4 degrees, as shown in Fig. 8.14, which is adequate in this application. In our upcoming report, we will present the measured characteristics of the on-axis HPM as well as experimental results on its performance as a receiver optical front-end. Fig. 8.15 shows a photograph of the on-axis holographic mirror.

(a)

(b)
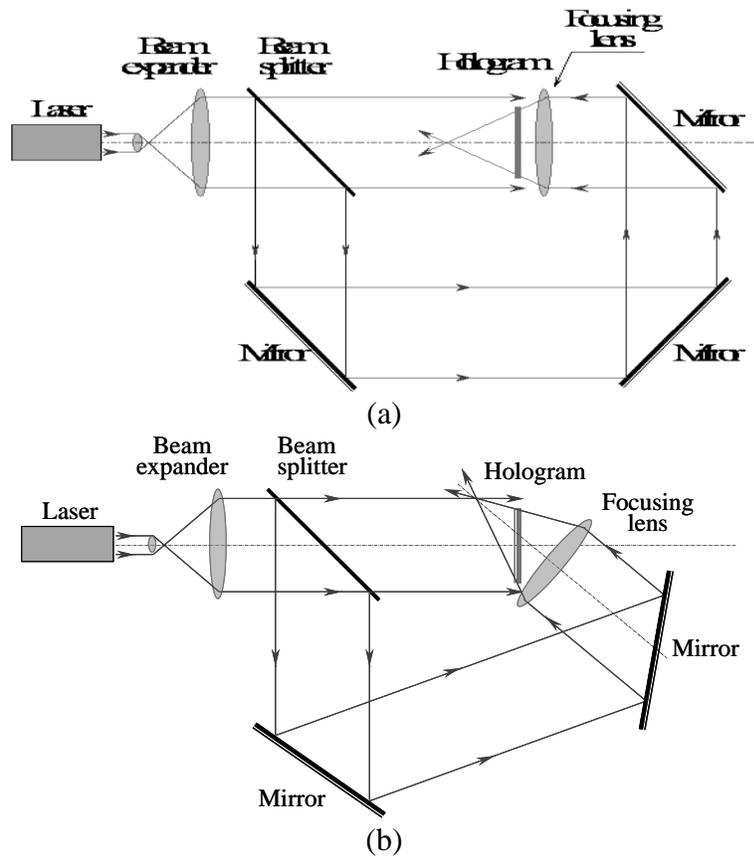
Fig. 8.13. Optical set-up for recording of (a) an on-axis parabolic holographic mirror, and (b) an off-axis parabolic holographic mirror.
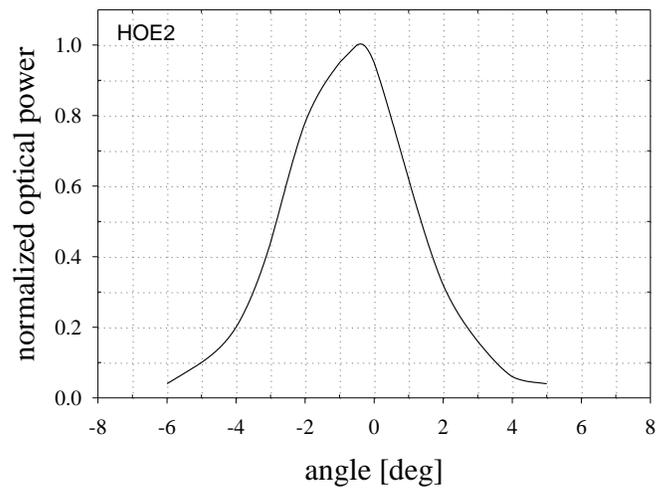
Fig. 8.14. Angular dependence of the optical power concentrated onto a photodiode positioned at the focal plane of an off-axis parabolic holographic mirror.

Fig. 8.15. Photograph of the recorded on-axis holographic parabolic mirror.

8.3.6 *Holographic Mirror and receiver Optical Subsystem Characteristics*

The holographic mirror characteristics (replay wavelength, spectral bandwidth, and diffraction efficiency) changed somewhat due to weak shrinkage after the initial swelling and stayed stable after that at their new values. No further change of these parameters has been observed until now (more than a month after the recording). The replay wavelength of the hologram shifted from 810 nm to 780 nm. Measurements using a laser diode with emission at 808 nm showed diffraction efficiency of only 2%, while a diffraction efficiency of 37% was measured using a laser diode with emission at 780 nm. The shift of the replay wavelength was confirmed by a measurement of the transmission spectrum of the hologram (Fig. 8.16). The spectral bandwidth of the mirror is 32 nm.
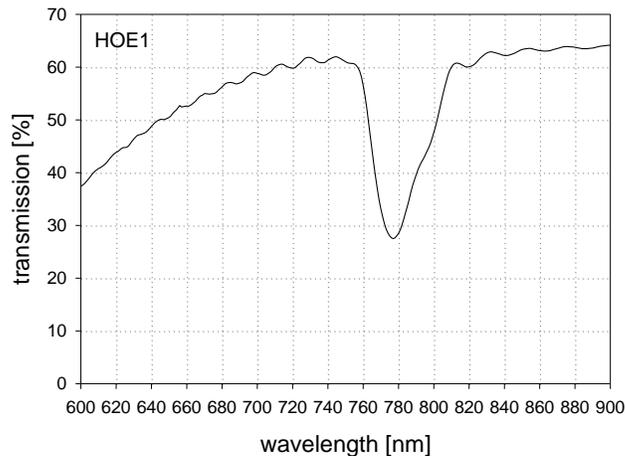
Fig. 8.16. Transmission spectrum of the on-axis HPM measured one year after recording.

The hologram diffraction efficiency and spectral bandwidth are not proper measures of the receiver optical front-end quality, though they give a rough idea about what to expect. The characteristics of the receiver optical front-end, such as field-of-view, concentration ratio, signal-to-noise ratio (SNR), etc., also depend on the characteristics of the hologram as an optical element. A hologram with very high diffraction efficiency can still be a bad concentrator if it exhibits strong optical aberrations. The photodiode characteristics

(responsivity, spectral bandwidth, etc.) as well as the size and the mutual position of the hologram and the photodiode have their impact on the receiver optical front-end characteristics, too. In our measurements a Si PIN photodiode was used with a sensitive area diameter of 6 mm and spectral responsivity bandwidth of about 500 nm (530 nm to 1030 nm). The photodiode was positioned in the focal plane of the holographic parabolic mirror and the angular response of the system mirror/photodiode was measured. The result is presented in Fig. 8.17.
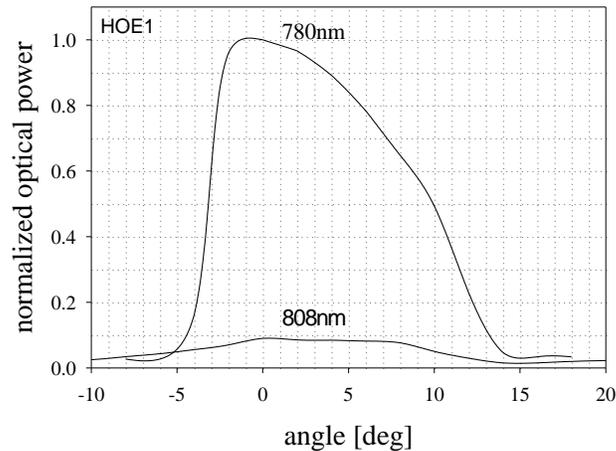
Fig. 8.17. Angular dependence of the optical power concentrated onto a photodiode positioned at the focal plane of an on-axis parabolic holographic mirror.

The angular characteristic of the holographic mirror is not symmetric with respect to the normal incidence. This may be caused by non-uniform swelling of the holographic emulsion. The full angle of acceptance at half height is 13 deg (FOV = 6.5 deg). Such a field-of-view can be classified as case 2, according to [80], in which at most one diffusing spot lies within a receiver branch field-of-view and practically unlimited channel frequency bandwidth is insured.

8.3.7   *Receiver Holographic Optical Front-End Performance*

The performance of the holographic mirror as a receiver optical front-end was evaluated. The receiver was placed about 88 cm away from a diffusely reflecting white screen, as shown in Fig. 8.18. An identical bare photodiode with unrestricted field-of-view was used as a reference receiver. A single diffusing spot on the screen within the field-of-view of the receiver was produced by a laser diode emitting 1.25 mW at 780 nm. A strong diffuse ambient light was simulated using the emission of a halogen lamp (Philips, 50 W).
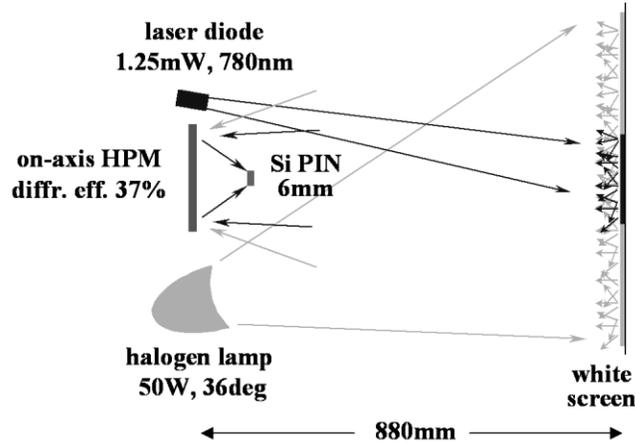
Fig. 8.18. A set-up to measure performance of the holographic mirror as a receiver optical front-end.

The measured signal gain, G_S, in comparison to a bare photodiode, was 6.5 dB at normal incidence. The measurement was performed with the laser diode on and the halogen lamp off. Significantly higher values can be obtained with a pure phase hologram (a bleached hologram or a hologram recorded in dichromated gelatin) having higher diffraction efficiency.

With the laser off and the lamp on, a very strong reduction in the ambient light reception was measured. The photodiode has maximum responsivity of 0.68 A/W at 940 nm and spectral bandwidth (responsivity at half height) of about 500 nm (530 nm to 1030 nm). The experimental background light gain, G_{bg}, (received ambient light compared to the amount of ambient light received by the reference bare photodiode) was -7.6 dB.

The signal gain, G_S, and the ambient light gain, G_{bg}, determine the improvement in the electrical signal-to-shot noise ratio (SNR) [3 - 4]; $G_{SNR} = \dfrac{G_s^2}{G_{bg}}$. Mainly, due to the very low value of G_{bg}, a tremendous improvement of 20.7 dB in the SNR was measured.

8.4 Conclusions

In conclusion, we can obtain transmit optics for the proposed active laser imaging system by employing efficient, low-cost photolithographic techniques to design holographic optical elements for beam-splitting. These can ensure the necessary transmit power and interference decoupling for the arrayed beamlet laser transmission that has been discussed in previous chapters.

On the receive side, an optical concentrator and ambient light rejection filter is implemented in the reflector-type holographic element. The presented study clearly shows that holographic mirrors are indeed very promising candidates as a receiver optical subsystem for infrared wireless communications. Particularly, an improvement in the diffraction efficiency of the hologram (e.g., employing a phase hologram) would increase the signal gain and the background light gain at the same time. However, the SNR is proportional to the square of the signal gain, and inversely proportional to only the first order of the background light gain. Thus, higher diffraction efficiency would improve the power efficiency of the receiver, as well as the signal-to-noise ratio.

Chapter 9
Conclusions

In this report, the problem of active optical imaging through cloudy turbulent atmosphere is investigated in depth and several mitigation techniques are explored. To this end, a comprehensive channel model is developed that takes the effect of both scattering and turbulence into account. A semi-analytical approach (Monte Carlo Ray Tracing) and a simplified Markov chain extracted from Monte Carlo are developed to model imaging through cloud, in conjunction with investigation of scintillation caused by atmospheric turbulence. Simulations were performed for active laser imaging with SISO sensing, with comparatively low optical thickness values. These initial investigations show that there is a possibility to perform active imaging in scattering and turbulent media, if shutter duration and field-of-view can be optimized. Furthermore, different geometric configurations (SISO mono-static and bi-static, and MIMO) have been tested. To realize optical imaging, a novel spatially-multiplexed MIMO optical imaging system, inspired by multi-spot diffuse indoor wireless optical communications is proposed to increase the area search rate, beam space loss, and hardware expenses. Using the framework developed in Chapters 2 to 5, performance of such a system under turbulent cloudy atmosphere is studied in Chapter 6 and a combination of several restoration modules used to reconstruct the distorted image is proposed in Chapter 7.

The preferred geometry is the MIMO configuration which uses a holographic beam splitter to illuminate the target in a pixel-wise manner. In this way, a uniform illumination of target is achieved. Furthermore, beam spreading is reduced by a factor of $1/N$, where N is the number of beamlets in each direction (i.e. we have an $N \times N$ array of beamlets). Moreover, due to smaller waists of these beamlets, the turbulence effects are significantly reduced. However, in design procedure of the photolithographic beam splitter, one should consider the following criteria: (a) Power efficiency of converting the Gaussian beam into array of beamlets, (b) Uniform distribution of intensity between the beamlets, (c) Minimization of interference (cross talk) between beamlets in far-field region. By using this multiplex MIMO configuration, the imaging problem is reduced to a point-to-point communications, i.e. a "to be, or not to be", scenario. In other words, performance metric is simplified to SNR per pixel. The viable resolution of this imaging system depends on the beam-splitter and focal plane array sizes and the atmospheric conditions. Atmospheric turbulence can cause interference between beamlets in far-field, and hence, gives rise to Inter-Pixel-Interference (IPI), which in turn reduces the resolution. As a result, the photolithographic beam-splitter must be designed properly.

To address the problem of backscatter, time-gating was investigated for the cloud dispersion model based on Monte-Carlo Ray Tracing. This method, though very simple, is shown to enhance the SNIR of the distorted image, thus improving the contrast of the image. A photo-detector system with fast shutters can be employed to implement the time-gating approach. We see improvements of 25 to 50dB over a non-time-gated method for different optical thickness values. An optimum value of the open shutter point and duration should be chosen so that most of the image information-bearing photons return at the same time as the least number of back-scattered photons. Good contrast may be achieved with the help of a post-processing method that applies an offline or near real-time time-gating on the laser reflected from the target.

Of particular interest will be to adaptively, in real time, steer the multiple transmitted beams (group beam forming) to achieve a better resolution of the image at the receive end. Since the transmitter and receiver are co-located, this very promising idea can be implemented.

Different methods of generating phase screens to emulate the effects of turbulence and scintillation are developed. Using this frame work, performance degradations in optical communications and imaging are clearly demonstrated and candidate solutions are proposed and evaluated both qualitatively and quantitatively. It was shown that RMDA family was more time-efficient, when it comes to generating a large ensemble of phase screens for atmospheric simulations. It is possible to extend these phase generation techniques to account for temporal variations in the atmospheric turbulence, based on "frozen" atmosphere assumption.

A combination of Adaptive Optics corrections and Blind Deconvolution is utilized to enhance the sharpness and resolution of the images. Three techniques of blind de-convolution were successfully implemented, under simulated conditions. Although this technique contains assumptions that may not be fulfilled by air-to-ground imaging systems, this lays the groundwork to obtain deconvolution of a true image from turbulence-induced PSF. In real-time, however, the algorithms are very much computationally demanding. The strength of additive noise also affects the convergence of the algorithm, and strong noise may hinder proper functioning. The supports of the true image and the blurring PSF are also required to be known, which are usually known or can be assumed. Acceleration of BD algorithm convergence after AO corrections is analyzed. With the help of these modules and provided that sufficient power resources are available, optical imaging through a turbulent, scattering medium is feasible.

The different fabrication technologies to obtain required transmit and receive optics are explored in Chapter 7. We reported on the fabrication of a 4x4 beam-splitter and its demonstration. A joint optimization of the branch field-of-view and the number of the diffusing spots on the image is necessary. Increasing the value of field-of-view means, increasing the level of ambient light detected by the receiver. Our investigations show that a photolithographic spherical mirror is a promising candidate as an optical front-end for the narrow field-of-view branches of a composite receiver used in a multi-spot configuration for optical imaging. Compared to a bare photo-detector, a holographic spherical mirror achieves 19.2dB improvement in the signal-to-noise ratio for normal incidence and 11.2dB improvement at the edge of the receiver branch field-of-view. Though other optical elements can exhibit a better performance, especially at oblique incidence, photolithographic spherical mirror has several important advantages: (a) multi-functionality (it combines concentrating and filtering functions), (b) insignificant physical weight, (c) small size, (d) low-cost when mass-produced. These factors play a significant role in optical imaging applications and can make the holographic curved mirror a preferable choice as a receiver optical subsystem. Despite the relatively low diffraction efficiency of holograms recorded in silver halide in principle, an improvement in the SNR of more than 20 dB has been measured. The use of photosensitive material within which a pure phase grating forms would significantly improve the signal gain of the receiver. Nevertheless, the results of the conducted experiments proved undoubtedly the great potential of curved photolithographic mirrors as a key element of the receiver optical front-end in laser imaging systems.

157

Appendix A
Assessment of Diversity on MIMO FSO links

This appendix summarizes the study that was undertaken to investigate the performance improvement on a MIMO lognormal FSO link, which is a simpler version of the imaging scenario discussed in the main body of the document. Although not directly related to imaging through turbulent atmosphere, the study sheds light on certain benefits that we can achieve on a MIMO free-space optic communication channel.

A.1   *MIMO System Model for FSO links*

In order to assess the performance of a multi-aperture setup for imaging, we first constructed a simplistic communications scenario. We assume one or multiple laser transmitters emitting towards multiple receivers. The transmitters may illuminate all of the receive apertures, or at least one of the receivers on its LOS path, and also illuminates the other neighboring receivers partially. The complete MIMO channel transfer matrix can be represented by

$$\mathbf{H}(\tau) = \begin{bmatrix} h_{11}(\tau) & \cdots & h_{1M}(\tau) \\ \vdots & \ddots & \vdots \\ h_{N1}(\tau) & \cdots & h_{NM}(\tau) \end{bmatrix} \tag{A.1}$$

where there are M transmitters, N receivers, and the channel response between the j-th transmitter and the i-th receiver is represented by $h_{ij}(\tau)$.

In a cloud-obscured optical channel, it can be expected that multiple scattering due to higher optical thickness increases the off-diagonal components, as a result of the spreading of the transmitted beams through cloud. However, the effect of clouds on the laser beam is deterministic at a macroscopic level, and the only randomness that will be manifested on the entries of $\mathbf{H}(\tau)$ is due to turbulence-induced fading. The fading coefficients are dependent on the receive aperture spacing and geometric configuration, and their variance could be reduced by having a large aperture so that aperture averaging holds, while not contributing as much MIMO gain.

In this section, we demonstrate through both analytical approximations and simulations, how a MIMO configuration could potentially improve the receive SNR of a diversity communications system. For now, we disregard the impulse responses due to cloud, and concentrate on finding possible performance improvements due to scintillation only, that may be relevant to imaging. The communications scenario is not directly related to imaging performance, but since the received total signal power would directly affect imaging performance (considering post-processing techniques), we can assess whether having a multi-aperture configuration benefits us significantly.

The limitation of these assessments is that we are not evaluating the distribution and variability of the received signal power on the detector surface points, such as evaluating SNR per pixel. This requires incorporation of point spread functions, mutual coherence functions, and optical transfer functions under scintillation, which will be discussed in later sections.

We consider a FSO link with M transmit and N receive apertures. We assume high SNR regime, where we can use Gaussian noise model. Assuming intensity-modulation/direct-

detection (IM/DD) and on-off keying (OOK) transmission [81], the received signal at the i-th receive aperture is given by

$$r_i = \eta A_r \sum_{j=1}^{M} h_{ij} s + v_i = a \sum_{j=1}^{M} h_{ij} s + v_i, i = 1, \ldots, N \quad (A.2)$$

where $s \in \left\{0, \sqrt{P_t}\right\}$ is the transmitted information bit, η is the optical-to-electrical conversion coefficient, A_r is the receive aperture area, and v_i is additive white Gaussian noise with zero mean and variance $\sigma_v^2 = N_o/2$. The fading channel coefficient which models the channel from the j-th transmit aperture to i-th receive aperture can be given based on the Rytov approximation,

$$h_{ij} = h_{0,ij} \exp\left(2\chi_{ij}\right) \quad (A.3)$$

where $h_{0,ij}$ is the signal light *intensity* without turbulence that shines from the j-th transmit aperture on the i-th receive aperture and χ_{ij} are identically distributed, but not necessarily independent, normal random variables with mean μ_χ and variance σ_χ^2. Therefore, h_{ij} follows a lognormal distribution

$$p(h_{ij}) = \frac{1}{2h_{ij}\sqrt{2\pi\sigma_\chi^2}} \exp\left(-\frac{\left(\ln(h_{ij}/h_{0,ij}) - 2\mu_\chi\right)^2}{8\sigma_\chi^2}\right) \quad (A.4)$$

To ensure that the fading does not attenuate or amplify the average power, we normalize the fading coefficients such that $E\left\{h_{ij}/h_{0,ij}\right\} = 1$. Doing so requires the choice of $\mu_\chi = -\sigma_\chi^2$. Assuming weak turbulence conditions, the variances of log-amplitude fluctuation of plane and spherical waves are given by [39],

$$\begin{aligned}
\sigma_{\chi,\mathrm{pl}}^2 &= 0.307 C_n^2 k^{7/6} L^{11/6} \\
\sigma_{\chi,\mathrm{sph}}^2 &= 0.124 C_n^2 k^{7/6} L^{11/6}
\end{aligned} \quad (A.5)$$

where $k = 2\pi/\lambda$ is the wavenumber, and L is the link distance in meters, C_n^2 is the refractive index structure coefficient and is altitude-dependent. The most commonly used C_n^2 profile model is the Hufnagle-Valley model [39],

$$C_n^2(h) = 0.00594(v/27)^2 (10^{-5}h)^{10} \exp(h/1000) + 2.7 \times 10^{-6} \exp(-h/1500) + A\exp(-h/1000) \quad (A.6)$$

where h is the altitude in meters, v is the rms wind speed in meters per second, and A is a nominal value of $C_n^2(0)$ at the ground in m$^{-2/3}$. For FSO links near the ground, C_n^2 can be taken approximately 1.7×10^{-14} m$^{-2/3}$ during daytime and 8.4×10^{-15} m$^{-2/3}$ at night. In general, C_n^2 varies from 10^{-13} m$^{-2/3}$ for strong turbulence to 10^{-17} m$^{-2/3}$ for weak turbulence with 10^{-15} m$^{-2/3}$ often defined as a typical average value.

Assuming $l_0 \leq \sqrt{\lambda L} \leq L_0$, where l_0 and L_0 are inner and outer scales, d_0, the correlation length of intensity fluctuations can be approximated by $d_0 \approx \sqrt{\lambda L}$. When the aperture size, $D_0 \gg d_0$, the detrimental effect of turbulence-induced fading is reduced due to aperture averaging. However, it is not always possible to make the aperture large enough, justifying the deployment of multiple photodetectors at the receiver side for scenarios with $D_0 < d_0$. The spatial correlation matrix **R** to model the correlations among receive apertures is given by

$$\mathbf{R} = \begin{bmatrix} 1 & b(d_{12}) & \cdots & b(d_{1N}) \\ b(d_{21}) & 1 & \cdots & b(d_{2N}) \\ \vdots & \vdots & \ddots & \vdots \\ b(d_{N1}) & b(d_{N2}) & \cdots & 1 \end{bmatrix} \tag{A.7}$$

where d_{ij} is the separation between i-th and j-th receiver apertures. $b(d)$ represents the normalized log-amplitude covariance function between two points in a receiving plane perpendicular to the direction of propagation and is defined by

$$b(d_{P_1 P_2}) = \frac{E[\chi(P_1)\chi(P_2)] - E[\chi(P_1)]E[\chi(P_2)]}{\sigma_\chi^2} \tag{A.8}$$

where $d_{P_1 P_2}$ is the distance between P_1 and P_2. Then the spatial covariance matrix $\mathbf{\Gamma}$ is given by $\mathbf{\Gamma} = \sigma_\chi^2 \mathbf{R}$. Similarly a correlation matrix of size $M \times M$ and corresponding covariance matrix can be defined for modeling the spatial correlations at the transmitter side.

A.2 Analysis & Simulation of Diversity Reception systems

In this section, we formulate SNR expressions for different diversity reception systems using multiple transmit and receive apertures. Closed-form approximations can be obtained for maximal ratio combining (MRC) and equal gain combining (EGC), using Schwarz and Yeh method of approximating sum of log-normal random variables. Simulations are also conducted and compared with analytical approximations.

A.2.1 Single Branch

With the assumptions given in Section 3, the received signal for a SISO system can be given by

$$r = a h_0 e^{2\chi} s + v \tag{A.9}$$

where $\chi \sim N(-\sigma_\chi^2, \sigma_\chi^2)$ and $v \sim N\left(0, \frac{N_o}{2}\right)$.

The SNR is easily seen to be

$$SNR_{SISO_\chi} = \frac{(a h_0 e^{2\chi})^2 E[s^2]}{E[v^2]} = \frac{(a h_0)^2 \cdot P_t/2}{N_o/2} e^{4\chi} = \frac{(a h_0)^2 P_t}{N_o} e^{4\chi} = SNR_0 e^{\hat{\chi}} \tag{A.10}$$

160

Therefore the SISO SNR is also a log-normal variable given by (A.10) where $\hat{\chi} \sim N\left(-4\sigma_\chi^2, 16\sigma_\chi^2\right)$ and SNR_0 is the SNR when no turbulence-induced fading is present.

A.2.2 MIMO with MRC

Next we consider a MIMO system with M transmit apertures and N receive apertures. For a fair comparison with the SISO system, the total transmitted intensity is assumed divided equally among the M transmitters (but can be divided unequally, given the total power constraint is met) and the total receiver aperture area is also kept the same by making each of the N receive apertures have an area of A_r/N. As a result, considering background noise limited receivers, the background noise for each receiver is also reduced by a factor of N.

Therefore, the signal received at the i-th receive aperture is given by

$$r_i = a \sum_{j=1}^{M} \frac{h_{0,ij}}{MN} e^{2\chi_{ij}} s + v_i \tag{A.11}$$

where $v_i \sim N\left(0, N_o/2N\right)$.

After maximal ratio combining, the decision metric can be written as

$$y = \sum_{i=1}^{N} \left(\left(a \sum_{j=1}^{M} \frac{h_{0,ij}}{MN} e^{2\chi_{ij}} \right)^2 s + \left(a \sum_{j=1}^{M} \frac{h_{0,ij}}{MN} e^{2\chi_{ij}} \right) v_i \right) \tag{A.12}$$

It must be noted that MRC requires proper channel tracking to know the fading coefficients.

The resulting SNR after MRC is

$$SNR_{MRC|\{\chi_{ij}\}} = \frac{E\left[\left(\sum_{i=1}^{N} \left(a \sum_{j=1}^{M} \frac{h_{0,ij}}{N} e^{2\chi_i} \right)^2 s \right)^2 \right]}{E\left[\left(\sum_{i=1}^{N} \left(a \sum_{j=1}^{M} \frac{h_{0,ij}}{MN} e^{2\chi_{ij}} \right) v_i \right)^2 \right]} = \frac{\left(\frac{ah_0}{MN} \right)^4 \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{M} e^{2\chi_{ij}} \right)^2 \right)^2 \cdot P_t/2}{\left(\frac{ah_0}{MN} \right)^2 \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{M} e^{2\chi_{ij}} \right)^2 \right) \cdot N_o/2N}$$

$$= \left(\frac{ah_0}{MN} \right)^2 \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{M} e^{2\chi_{ij}} \right)^2 \right) \cdot \frac{P_t}{N_o} N = \frac{1}{M^2 N} \cdot \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{M} e^{2\chi_{ij}} \right)^2 \right) \cdot SNR_0 \tag{A.13}$$

For the above formulation, the transmitted intensity is assumed equally divided among different transmitters, i.e. $h_{0,ij} = \dfrac{h_0}{M}$. But the calculation can be redone with different channel gains on the different transmission paths.

A.2.2.1 Approximation of Sum of Correlated Log-normal Variables

We can use Schwarz and Yeh method to find the mean and variance of the lognormal variable that approximates the sum of squares of a sum of lognormal variables in the above equation. This is based on the following approximation technique for the summation of

(correlated) log-normal random variables. Specifically, we want to approximate $\sum_{k=1}^{K} \exp(u_k)$ with a single log-normal random variable e^z where z is a Gaussian random variable with a mean of μ_z and a variance of σ_z^2. Defining $\text{cov}(u_k, u_l) = v_{kl}$, $E(u_k) = \mu_k$ and using results derived by Schwarz and Yeh, we have

$$\mu_z = \log\left(\alpha / \sqrt{1 + \beta^2/\alpha^2}\right), \quad \sigma_z^2 = \log\left(1 + \beta^2/\alpha^2\right) \tag{A.14}$$

Here α and β are defined as

$$\alpha = \sum_{k=1}^{K} e^{\mu_k + v_{kk}/2}, \quad \beta^2 = \sum_{k=1}^{K}\sum_{l=1}^{K} e^{\mu_k + \mu_l + \frac{(v_{kk} + v_{ll})}{2}} \left(e^{v_{kl}} - 1\right) \tag{A.15}$$

Note that, due to power normalization (i.e. $E[I/I_0] = E[\exp(2\chi_k)] = E[\exp(u_k)] = 1$), we have $E(\chi_k) = -\sigma_\chi^2$ and $\mu_k = -v_{kk}/2$. Further, under the assumption of weak turbulence (i.e., $v_{kk} \ll 1$), and small correlation values (i.e., $v_{kl} \ll 1$), μ_z and σ_z^2 can be simplified as

$$\mu_z = \log K - \sigma_z^2/2, \quad \sigma_z^2 \approx \frac{1}{K^2}\sum_{k,l} e^{v_{kl}} - 1 \approx \frac{1}{K}v_{11} + \frac{1}{K^2}\sum_{k \neq l} v_{kl} \tag{A.16}$$

where we assume that log-normal parameters are equal for all channels, i.e., $v_{kk} = v_{11} = 4\sigma_\chi^2$.

A.2.3 MIMO with EGC

With similar MIMO setup as for MRC, the output of an equal gain combiner can be written as

$$y = s\sum_{i=1}^{N}\sum_{j=1}^{M} \frac{ah_0}{MN} e^{2\chi_{ij}} + \sum_{i=1}^{N} v_i \tag{A.17}$$

The SNR at the output of the combiner, then, is represented as

$$\begin{aligned}
SNR_{EGC|\{\chi_{ij}\}} &= \frac{E\left[\left(s\sum_{i=1}^{N}\sum_{j=1}^{M} \frac{ah_0}{MN} e^{2\chi_{ij}}\right)^2\right]}{E\left[\left(\sum_{i=1}^{N} v_i\right)^2\right]} = \frac{P_t \Big/ 2 \cdot \left(\frac{ah_0}{MN}\right)^2 \cdot \left(\sum_{i=1}^{N}\sum_{j=1}^{M} e^{2\chi_i}\right)^2}{N_o/2} \\[2ex]
&= \left(\frac{ah_0}{MN}\right)^2 \left(\sum_{i=1}^{N}\sum_{j=1}^{M} e^{2\chi_i}\right)^2 \frac{P_t}{N_o} = \left(\frac{1}{MN}\right)^2 \left(\sum_{i=1}^{N}\sum_{j=1}^{M} e^{2\chi_i}\right)^2 \cdot SNR_0
\end{aligned} \tag{A.18}$$

A.3   Diversity Combining Results and Conclusions

Based on the formulations in Sections A.1 and A.2, two different turbulence conditions with $\sigma_x = 0.1$ (weak) and $\sigma_x = 0.3$ (strong) were considered. The plots in Fig. A.1 and Fig. A.2 show simulated and analytical (approximated) probability density functions of the output SNRs of the different combining schemes for both uncorrelated fading and correlated fading with correlation coefficients of $\rho = \{0, 0.1, 0.3, 0.9\}$.

(a) $\rho = 0$

(b) $\rho = 0.1$

(c) $\rho = 0.3$

(d) $\rho = 0.4$
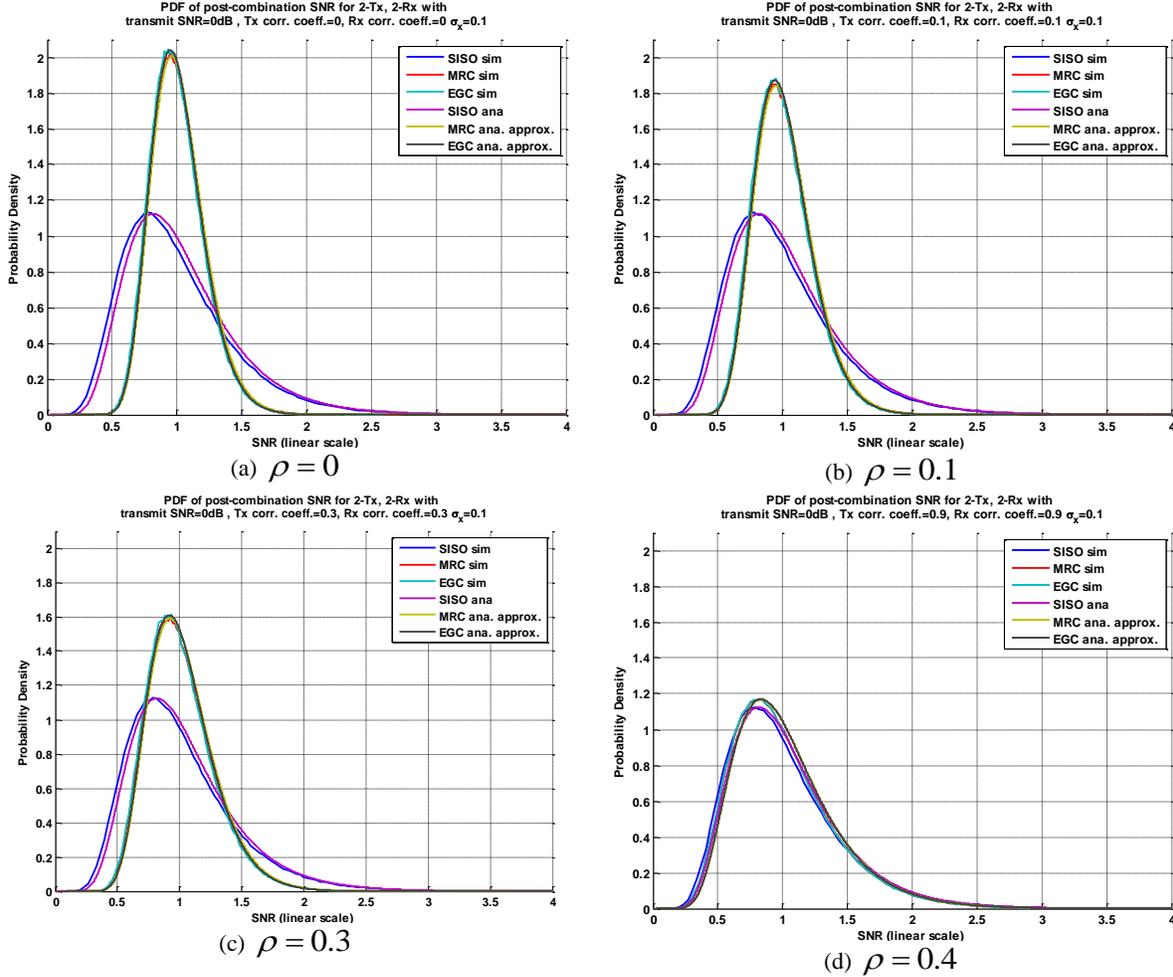
Fig. A.1 (a-d). PDF of post-detection SNR for a 2-Tx 2-Rx MIMO FSO system with $SNR_0 = 1$, and log-normal fading variance, $\sigma_\chi^2 = 0.01$, correlation coefficient, $\rho = \{0, 0.1, 0.3, 0.9\}$. All in linear scale.

PDF of post-combination SNR for 2-Tx, 2-Rx with
transmit SNR=0dB , Tx corr. coeff.=0, Rx corr. coeff.=0 σ_x=0.3

(a) $\rho = 0$

PDF of post-combination SNR for 2-Tx, 2-Rx with
transmit SNR=0dB , Tx corr. coeff.=0.1, Rx corr. coeff.=0.1 σ_x=0.3

(b) $\rho = 0.1$

PDF of post-combination SNR for 2-Tx, 2-Rx with
transmit SNR=0dB , Tx corr. coeff.=0.3, Rx corr. coeff.=0.3 σ_x=0.3

(c) $\rho = 0.3$

PDF of post-combination SNR for 2-Tx, 2-Rx with
transmit SNR=0dB , Tx corr. coeff.=0.9, Rx corr. coeff.=0.9 σ_x=0.3

(d) $\rho = 0.4$

Fig. A.2 (a-d). PDF of post-detection SNR for a 2-Tx 2-Rx MIMO FSO system with $SNR_0 = 1$, and log-normal fading variance, $\sigma_\chi^2 = 0.09$, correlation coefficient, $\rho = \{0, 0.1, 0.3, 0.9\}$. All in log scale.

It is evident that in all cases MIMO shifts the SNR PDFs further to the right from the SISO case, and reduces the variability of the SNR. This means that, in a diversity-based multiple-aperture system, under the given turbulence conditions with given correlations, will perform better on the average than a SISO system.

It is also observed that a large correlation coefficient performs at an intermediate point between fully uncorrelated MIMO diversity system and a SISO system. The covariance values of $\sigma_x = 0.1$, 0.3 and correlation coefficients $\rho = 0$ through 0.9, are based on the work by [5][86].

One interesting observation is that MRC does not perform better than EGC. This is counter-intuitive to results from RF communication systems. However, here we have considered intensity-modulation and direct detection, due to which the current output of the photodetector is proportional to the power, and maximal ratio combining results in noise being enhanced, and thus SNR reduced, while EGC results in equal-gain addition and noise is not amplified. The results indicate this as a fact for IM/DD.

In the context of an imaging system, this diversity mechanism corresponds to the scenario when the multiple transmitters shine the laser beam on the same object with identical footprints, and the multiple receive apertures are all directed at the same cross-section of the object. This does not provide multiple viewing directions on the object, but would provide better signal to noise ratios per pixel, according to the results plotted in Fig. A.1 and Fig. A.2. The assumptions to arrive at these results, however, should be kept in mind, considering practicality issues. The two principal assumptions in arriving at these results are (1) $d_0 \approx \sqrt{\lambda L}$, the correlation length can be approximated by the first Fresnel zone, which is ~4cm for a 1.55-µm laser for a propagation distance of 1km, and (2) $D_0 < d_0$, the aperture size is smaller than the correlation length. Furthermore, the Rytov approximation and adaptive optics, as necessary for spatial phase conjugation, is assumed. The link is assumed horizontal in contrast to a vertical link that is relevant for imaging. These assumptions, therefore, have to be investigated or modified for an imaging scenario. In reality, these assumptions may be too stringent, or may result in other complications related to imaging, such as, possible degradation of image resolution with increasing lens diameter, which warrants further investigations into the statistical profiles of point spread functions, which are discussed in later sections.

A.4   MIMO Diversity Order Assessment for FSO links

The benefit of using multi-aperture configuration in a log-normal fading environment can be twofold. Either we can achieve a diversity gain, as seen in Sections A.2 through A.3, whereby we improve the received SNR and obtain better probability of outage for a given data rate; or, we can achieve a multiplexing gain, through which we can increase the data transmission capacity of a communication link or the area search rate in an imaging system, while keeping the SNR and outage probabilities similar to a SISO link.

We concentrate on the diversity scenario to quantify how a MIMO configuration can improve system performance parameters, such as SNR and outage probability. In terms of an optical imaging system, this quantification is not entirely appropriate; however, since the total image power and the availability are the parameters that will most directly limit the performance of an optical imaging system, this assessment would indicate whether a multiple transmit/receive aperture system would be beneficial for imaging systems. Diversity system essentially means that the same cross-section of a target object is illuminated by the array of imaging transmitters and the same cross-section is focused on by the receiving array.

Spatial diversity is an effective tool to mitigate the degrading effects of fading and has been extensively studied by several authors [81][82][83][84] within the context of FSO communication. The turbulence-induced fading model commonly adopted in FSO communications is log-normal distribution. In this study, we adopt relative diversity order (RDO) [81], as a performance measure for benchmarking MIMO FSO channels against SISO channels. For a MIMO diversity system, there are several means of combining power/information received by multiple receivers, namely (i) Selection Combining, (ii) Maximal Ratio Combining, (iii) Equal Gain Combining, etc. The last two are of our main interest in this study, since MRC is the optimal scheme, and EGC is a sub-optimal scheme, which sometimes approaches MRC under certain conditions, as seen in Section A.2. Both of these schemes can be implemented for a FSO system, even though they may not be as simple as a SC diversity system.

The assumptions for the RDO assessment study are as follows [81][85]. A FSO IM-DD system employing binary pulse position modulation is considered, where the optical transmitter is "on" during half duration of the bit interval and is "off" during the other half. The receivers integrate both the signal and non-signal slots of the pulse and we obtain the resulting signal vector given by

$$\mathbf{r} = \begin{bmatrix} r^s & r^n \end{bmatrix}^T = \begin{bmatrix} \eta T (P_s + P_b) + n^s & \eta T P_b + n^n \end{bmatrix}^T \tag{A.19}$$

Here, r^s and r^n are the received electrical signals which correspond to the signal and non-signal slots of the BPPM pulse. P_s and P_b are the optical signal power and the background power incident on the photo-detector, T is the bit duration, η is the responsivity of the photo-detector(s), and n^s and n^n are the additive white Gaussian noise terms in "on" and "off" slots with zero-mean and variance $\sigma_n^2 = N_o/2$. The receivers are also assumed background noise-limited [5][86].

In the MIMO FSO system, the source transmits an intensity-modulated signal through M transmit apertures and the destination node receives the faded signal using direct detection through N receive apertures. The destination node combines the outputs of the receive apertures before decision-making.

Incorporating distance-dependent path loss and turbulence-induced fading, the channel gain of a link of length d can be given as, $g = |\alpha|^2 PL(d)$, where $|\alpha| = \exp(\chi)$ is the log-normally distributed channel fading *amplitude* which is commonly used to model *weak* turbulence conditions. The fading log-amplitude χ is normally distributed with mean μ_χ, and variance σ_χ^2. To ensure that fading does not attenuate or amplify the average power, the fading amplitude is normalized so that $E\left[|\alpha|^2\right] = 1$, which is possible by making $\mu_\chi = -\sigma_\chi^2$. The term $PL(d) = l(d)/l(d_{s,d})$ denotes the normalized path loss with respect to the distance of the direct link between the source and the destination, $d_{s,d}$. For a long transmission link, a suitable assumption is that the distances between individual transmitters and receivers are approximately equal to $d_{s,d}$; and thus the normalized path loss term can be assumed to be 1 [82].

Let us denote an instantaneous electrical SNR γ corresponding to an instantaneous channel realization $\alpha = \alpha'$. The outage probability at the transmission rate R_0 is given by $P_{out}(R_0) = \Pr\{C(\gamma) < R_0\}$, where $C(\gamma)$ is the instantaneous capacity corresponding to the instantaneous SNR γ. Since $C(\gamma)$ is a monotonically increasing function of γ, the outage probability can also be expressed in terms of SNR as [81],

$$P_{out}(R_0) = \Pr\{\gamma < \gamma_{th}\} \tag{A.20}$$

where $\gamma_{th} = C^{-1}(R_0)$ is the threshold SNR. If SNR exceeds γ_{th}, no outage happens and signal can be decoded with arbitrarily low error probability at the receiver.

Based on the above model SNR and outage probability functions can be formulated as follows for different diversity schemes, including SISO.

A.4.1   SISO

The optical signal power incident on the photodetector at the destination can be expressed as $P_s = P_t g_{s,d} = P_t |\alpha_{s,d}|^2$, where $g_{s,d}$ and $|\alpha_{s,d}|$ are the gain and fading amplitude of the channel linking the source and destination nodes and P_t is the total average transmit power. Thus, the received electrical SNR at the destination node can be obtained as

$$\gamma = \frac{\eta^2 T^2 P_s^2}{N_o} = \frac{\eta^2 T^2 P_s^2 |\alpha_{s,d}|^4}{N_o} \tag{A.21}$$

The outage probability given by (6) is then,

$$P_{out,SISO} = \Pr\{\gamma < \gamma_{th}\} = \Pr\left\{\frac{\eta^2 T^2 P_t^2 |\alpha_{s,d}|^4}{N_o} < \gamma_{th}\right\} = \Pr\left\{|\alpha_{s,d}|^4 < \frac{\gamma_{th} N_o}{\eta^2 T^2 P_t^2}\right\}$$

$$= \Pr\left\{|\alpha_{s,d}|^2 < \sqrt{\frac{\gamma_{th} N_o}{\eta^2 T^2 P_t^2}}\right\} = \Pr\left\{|\alpha_{s,d}|^2 < 1/P_M\right\} \tag{A.22}$$

where P_M denotes the power margin and is defined as $P_M = P_t / P_{th}$. P_{th} denotes a threshold transmit power required to guarantee that no outage happens in a direct fading-free transmission from the source to the destination. Thus, the power margin can be expressed as $P_M = \sqrt{\eta^2 P_t^2 T^2 / N_o \gamma_{th}}$. The outage probability of (A.22) can also be expressed in terms of the log-normal Cumulative Distribution Function (CDF) as follows,

$$P_{out,SISO}(P_M) = CDF_{\text{log-norm}}\left(1/P_M, -2\sigma_\chi^2, 4\sigma_\chi^2\right) \tag{A.23}$$

where $CDF_{\text{log-norm}}\left(x, \mu, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{e^{\frac{-(\ln t - \mu)^2}{2\sigma^2}}}{t} \cdot dt$.

For simulation purposes, a long series of lognormal random variables χ are generated with mean $\mu = -\sigma_\chi^2$ and variance of σ_χ^2, from which the random variables $\alpha_{s,d}$ are determined, and then the outage probability is calculated by evaluating the probability given in (A.22) for different values of P_M.

A.4.2   MIMO MRC

The optical signal power incident on the photodetector at i-th receive aperture of the destination node can be expressed as $P_s(i) = (P/N) \cdot \sum_{j=1}^M g_{i,j} = (P/N) \cdot \sum_{j=1}^M |\alpha_{i,j}|^2$, where $P = P_t/M$ is the average transmitted optical power per transmit aperture. To make a fair comparison, we assume that the sum of the N receive aperture areas is the same as the area of the receive aperture for a SISO link. For maximal ratio combining the received electrical signals on individual receiving apertures are multiplied by their conjugates and added together before decoding. Since this is an IM-DD system, the received power does not have any phase, and the received electrical current, which is proportional to optical power, is simply squared,

before addition. Furthermore, the background noise at the receiving apertures is also reduced by a factor of N. Therefore, the received SNR for the MIMO-MRC scheme can be written as

$$\gamma = \frac{\eta^2 T^2 P_t^2}{M^2 N^2 \cdot N_o/N} \cdot \sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2 = \frac{\eta^2 T^2 P_t^2 / M^2 N \cdot \sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2}{N_o} \quad\quad \text{(A.24)}$$

The outage probability defined by (A.22) for MIMO-MRC then becomes,

$$
\begin{aligned}
P_{out,MRC} = \Pr\{\gamma < \gamma_{th}\} &= \Pr\left\{\frac{\eta^2 T^2 P_t^2 / M^2 N \cdot \sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2}{N_o} < \gamma_{th}\right\} \\[2mm]
&= \Pr\left\{\frac{1}{M^2 N}\sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2 < \frac{\gamma_{th}N_o}{\eta^2 T^2 P_t^2}\right\} \\[2mm]
&= \Pr\left\{\frac{1}{M^2 N}\sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2 < 1/P_M^2\right\}
\end{aligned}
\quad\quad \text{(A.25)}
$$

where similar definitions as before hold.

Nothing has been mentioned so far about the correlations between the fading log-normal variables. The correlation among the log-normal fading coefficients is described by the transmit-side and receive-side correlation matrices R_{Tx} and R_{Rx} [82]. The i,j-th entry of these matrices, $\rho_{Tx,ij}$ or $\rho_{Rx,ij}$, indicate the normalized correlation coefficient between the i-th and j-th transmitters or receivers. For simplicity, it is assumed that the geometric configurations of the transmitter arrays are such that $\rho_{Tx,ij} = \rho_{Tx}, i \neq j$ and $\rho_{Tx,ii} = 1, \forall i$, and similar for the receiver arrays. The normalized full correlation matrix describing correlation statistics among all transmitters and receivers is obtained from the Kronecker product of R_{Tx} and R_{Rx}, i.e., $R = R_{Tx} \otimes R_{Rx}$. The normalization can be removed by multiplying R by σ_χ^2.

By employing approximation methods based on works by Schwartz and Yeh [82], the sum of log-normal variables $\sum_{i=1}^{N}\left(\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2$ can be approximated in the following manner. We can represent the sum inside the first squared term as a single log-normal variable, $u_{1,i}$, $\sum_{j=1}^{M}|\alpha_{i,j}|^2 = \sum_{j=1}^{M}e^{2\chi} \approx e^{u_{1,i}}$. Variable $u_{1,i}$ will be log-normal with mean and variance given by,

$\mu_{u_1} = \log\left(\alpha_1/\sqrt{1+\beta_1^2/\alpha_1^2}\right)$ and $\sigma_{u_1}^2 = \log\left(1+\beta_1^2/\alpha_1^2\right)$, where $\alpha_1 = \sum_{k=1}^{M}e^{\mu_{2\chi}+\sigma_{2\chi}^2/2} = \sum_{k=1}^{M}e^{-2\sigma_\chi^2+4\sigma_\chi^2/2} = M$ and

$\beta_1^2 = \sum_{k=1}^{M}\sum_{l=1}^{M}e^{2\mu_{2\chi}+\sigma_{2\chi}^2}\left(e^{\rho_{Tx,kl}\sigma_{2\chi}^2}-1\right) = M\left(e^{4\sigma_\chi^2}-1\right)+M(M-1)\left(e^{4\rho_{Tx}\sigma_\chi^2}-1\right)$. The outer sum $\sum_{i=1}^{N}e^{2u_{1,i}}$ can then be approximated as another log-normal variable u_2, i.e. $\sum_{i=1}^{N}e^{2u_{1,i}} = e^{u_2}$. Variable u_2 is then another

log-normal variable with mean and variance given by, $\mu_{u_2} = \log\left(\alpha_2 \big/ \sqrt{1 + \beta_2^2 / \alpha_2^2}\right)$ and

$$\sigma_{u_2}^2 = \log\left(1 + \beta_2^2 / \alpha_2^2\right), \qquad \text{where} \qquad \alpha_2 = \sum_{k=1}^{N} e^{\mu_{2u_1} + \sigma_{2u_1}^2 / 2} = N \cdot e^{2\left(\mu_{u_1} + \sigma_{u_1}^2\right)}, \qquad \text{and}$$

$$\beta_2^2 = \sum_{k=1}^{N} \sum_{l=1}^{N} e^{2\mu_{2u_1} + \sigma_{2u_1}^2} \left(e^{\rho_{Rx,kl}\sigma_{2u_1}^2} - 1\right) = \left[e^{4\left(\mu_{u_1} + \sigma_{u_1}^2\right)}\right]\left[N\left(e^{4\sigma_{u_1}^2} - 1\right) + N(N-1)\left(e^{4\rho_{Rx}\sigma_{u_1}^2} - 1\right)\right].$$

Thus, the outage probability of (A.25) can also be expressed in terms of the log-normal Cumulative Distribution Function (CDF) as follows,

$$P_{out,MRC}\left(P_M\right) = CDF_{\text{log-norm}}\left(1 / P_M^2, \mu_{u_2} - \log\left(M^2 N\right), \sigma_{u_2}^2\right) \tag{A.26}$$

For simulation purposes, a long series of lognormal random variables χ are generated with mean $\mu = -\sigma_\chi^2$ and variance of σ_χ^2, and they are filtered by $R^{1/2}$, the matrix square root of the full correlation matrix (unnormalized). Afterwards, their mean is shifted to the one desired before the random variables $\alpha_{i,j}$ are determined. Then, the above outage probability is calculated by evaluating the probability given in (A.25), for different values of P_M.

A.4.3 MIMO EGC

Similar as for MIMO-MRC, the optical signal power incident on the photodetector at i-th receive aperture of the destination node can be expressed as $P_s(i) = P/N \cdot \sum_{j=1}^{M} g_{i,j} = P/N \cdot \sum_{j=1}^{M} |\alpha_{i,j}|^2$, where $P = P_t/M$ is the average transmitted optical power per transmit aperture. To make a fair comparison, we again assume that the sum of the N receive aperture areas is the same as the area of the receive aperture for a SISO link [5].

For equal gain combining, the received electrical signals on individual receiving apertures are phase-conjugated and added together before decoding. Since this is an IM-DD system, the received power does not have any phase, and the electrical current, which is proportional to optical power, is simply added. Therefore, the received SNR for the MIMO-EGC scheme can be written as

$$\gamma = \frac{\eta^2 T^2 P_t^2}{M^2 N^2 \cdot N_o}\left(\sum_{i=1}^{N}\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2 = \frac{\eta^2 T^2 P_t^2 / M^2 N^2 \cdot \left(\sum_{i=1}^{N}\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2}{N_o} \tag{A.27}$$

The outage probability of (A.22) then can be given by,

$$
\begin{aligned}
P_{out,EGC} &= \Pr\left\{\gamma < \gamma_{th}\right\} = \Pr\left\{\frac{\eta^2 T^2 P_t^2 / M^2 N^2 \cdot \left(\sum_{i=1}^{N}\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2}{N_o} < \gamma_{th}\right\} \\
&= \Pr\left\{\frac{1}{M^2 N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{M}|\alpha_{i,j}|^2\right)^2 < \frac{\gamma_{th} N_o}{\eta^2 T^2 P_t^2}\right\} = \Pr\left\{\frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}|\alpha_{i,j}|^2 < 1/P_M\right\}
\end{aligned}
\tag{A.28}
$$

where similar definitions as before hold.

The correlation among the log-normal fading coefficients is similarly defined as for MIMO-MRC.

Again, by employing approximation methods for sum of log-normal random variables, the inner sum $\sum_{j=1}^{M}|\alpha_{i,j}|^2$ of (A.28) can be represented as a single log-normal variable $u_{1,i}$, i.e. $\sum_{j=1}^{M}|\alpha_{i,j}|^2 = \sum_{j=1}^{M}e^{2\chi} \approx e^{u_{1,i}}$. Variable $u_{1,i}$ will be log-normal with mean and variance given by,

$\mu_{u_1} = \log\left(\alpha_1 / \sqrt{1 + \beta_1^2/\alpha_1^2}\right)$ and $\sigma_{u_1}^2 = \log\left(1 + \beta_1^2/\alpha_1^2\right)$, where $\alpha_1 = \sum_{k=1}^{M}e^{\mu_{2\chi} + \sigma_{2\chi}^2/2} = \sum_{k=1}^{M}e^{-2\sigma_\chi^2 + 4\sigma_\chi^2/2} = M$ and

$\beta_1^2 = \sum_{k=1}^{M}\sum_{l=1}^{M}e^{2\mu_{2\chi} + \sigma_{2\chi}^2}\left(e^{\rho_{Tx,kl}\sigma_{2\chi}^2} - 1\right) = M\left(e^{4\sigma_\chi^2} - 1\right) + M(M-1)\left(e^{4\rho_{Tx}\sigma_\chi^2} - 1\right)$. The outer sum $\sum_{i=1}^{N}e^{u_{1,i}}$ can then

be approximated as another log-normal variable u_2, i.e. $\sum_{i=1}^{N}e^{u_{1,i}} = e^{u_2}$. Variable u_2 is then another log-normal variable with mean and variance given by, $\mu_{u_2} = \log\left(\alpha_2 / \sqrt{1 + \beta_2^2/\alpha_2^2}\right)$ and

$\sigma_{u_2}^2 = \log\left(1 + \beta_2^2/\alpha_2^2\right)$, where $\alpha_2 = \sum_{k=1}^{N}e^{\mu_{u_1} + \sigma_{u_1}^2/2} = MN$, and

$\beta_2^2 = \sum_{k=1}^{N}\sum_{l=1}^{N}e^{2\mu_{u_1} + \sigma_{u_1}^2}\left(e^{\rho_{Rx,kl}\sigma_{u_1}^2} - 1\right) = M^2\left[N\left(e^{\sigma_{u_1}^2} - 1\right) + M(M-1)\left(e^{\rho_{Rx}\sigma_{u_1}^2} - 1\right)\right]$.

Thus, the outage probability of (A.28) can also be expressed in terms of the log-normal Cumulative Distribution Function (CDF) as follows,

$$P_{out,EGC}(P_M) = CDF_{\text{log-norm}}\left(1/P_M, \mu_{u_2} - \log(MN), \sigma_{u_2}^2\right) \tag{A.29}$$

For simulation purposes, multiple log-normal variables are generated similarly as for MIMO-MRC case, and filtered to introduce correlation. Then the probability of outage is evaluated according to (A.28).

A.5   *Results for Diversity Order Assessment and Conclusions*

Diversity order is defined as the negative of the asymptotic slope of the error rate performance (e.g. bit error rate or outage probability) versus SNR. With this convention, the diversity order of a SISO transmission is given by

$$d = -\lim_{SNR \to \infty} \frac{\partial \ln P_{out}}{\partial \ln SNR} = -\lim_{SNR \to \infty} \frac{\partial \ln P_{out}}{\partial \ln P_M} \tag{A.30}$$

where P_{out} is the outage probability, and P_M is the power margin.

Using expressions of outage probability obtained by [81], it can be shown that the conventional diversity order tends to infinity for log-normally faded FSO channels. It is due to this reason that *relative* diversity order measure is necessary. The single-branch diversity system is taken as the benchmark, and then the relative diversity order (RDO) is quantified as [81],

$$RDO(P_M) = \frac{\partial \ln P_{out}/\partial \ln P_M}{\partial \ln P_{out,SISO}/\partial \ln P_M} \tag{A.31}$$

Furthermore, the asymptotic relative diversity order (ARDO) can be given by

$$ARDO = \lim_{P_M \to \infty} RDO(P_M) \tag{A.32}$$

The RDO and ARDO quantities can be determined by numerically differentiating the probability of outage, P_{out}, as calculated in Section A.4, with respect to P_M, the power margin.

This section presents simulation and analytical approximation results for the RDO and ARDO of different MIMO systems. Overall, two conditions are considered, (i) weak scintillation represented by log-normal variance, $\sigma_\chi^2 = 0.01$, and (ii) strong scintillation represented by log-normal variance, $\sigma_\chi^2 = 0.09$. The correlation coefficients on transmit and receive sides are kept the same, and the values considered are, $\rho_{Tx} = \rho_{Rx} = \{0, 0.1, 0.3, 0.9\}$.

$$\rho_{Tx} = \rho_{Rx} = 0 \qquad\qquad \rho_{Tx} = \rho_{Rx} = 0.1$$

$$\rho_{Tx} = \rho_{Rx} = 0.3 \qquad\qquad \rho_{Tx} = \rho_{Rx} = 0.9$$
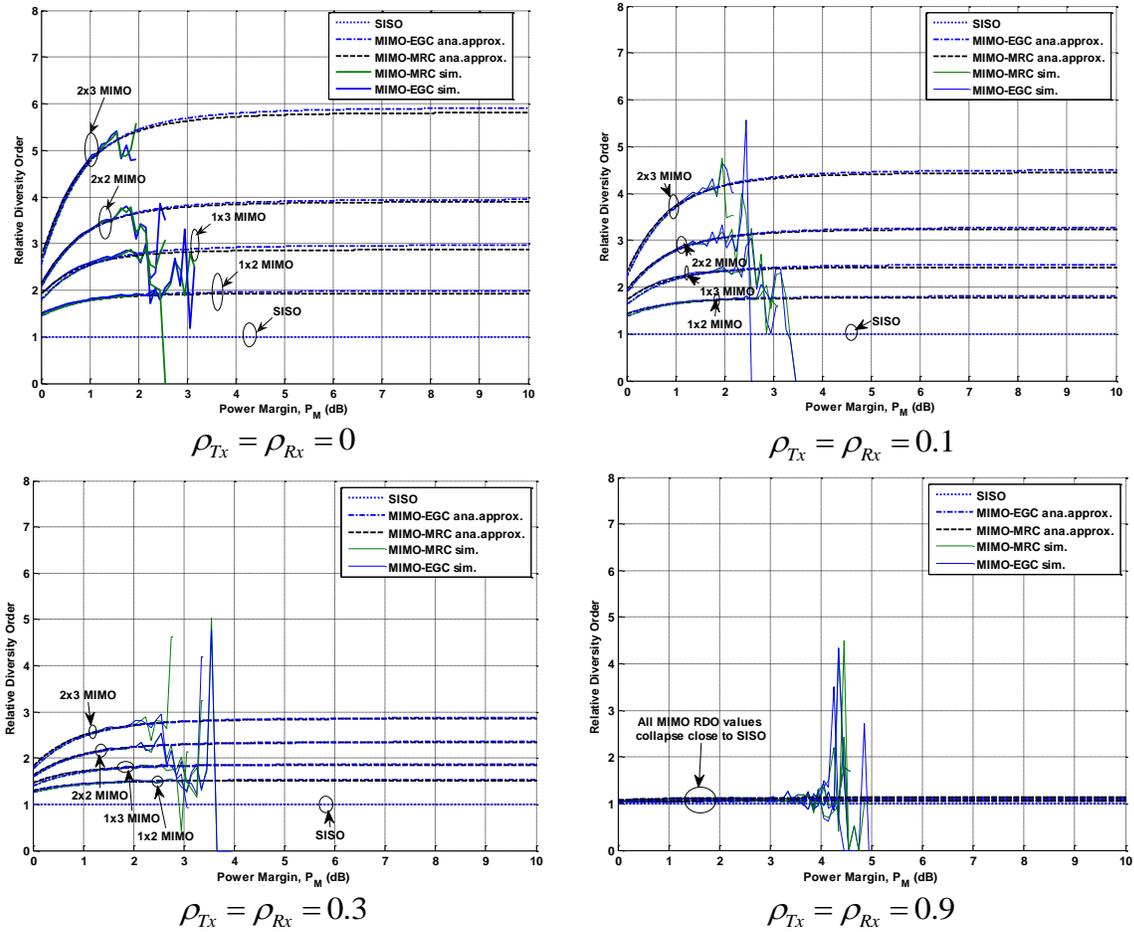
Fig. A.3 Relative Diversity Order versus Power Margin in dB for
$$\sigma_\chi^2 = 0.01, \rho_{Tx} = \rho_{Rx} = \{0, 0.1, 0.3, 0.9\}$$

Fig. A.3 shows the relative diversity order values plotted for the above-mentioned correlation coefficient values for a log-normal variance of $\sigma_\chi^2 = 0.01$. The diversity order achieved by MIMO increases with increasing number of transmitters and receivers. The

Asymptotic Diversity Order can be approximated by the value of the flattening line towards infinity, and the ARDO is seen to vary significantly with correlation values, from a maximum of *MN* (for full correlation, $\rho_{Tx} = \rho_{Rx} = 0$) to a minimum of 1 (for full correlation, $\rho_{Tx} = \rho_{Rx} = 1$). The analytical approximations seem to fit quite well with the simulated values. Neither the simulated results, nor the analytical approximations show any difference between the diversity order obtained by employing MRC and EGC. The oscillations caused at the tail of the simulated values should be discarded, since about 10^8 realizations were used for the simulations, and this value was not adequate to get enough points to have a smooth tail probability.

$$\rho_{Tx} = \rho_{Rx} = 0 \qquad\qquad \rho_{Tx} = \rho_{Rx} = 0.1$$

$$\rho_{Tx} = \rho_{Rx} = 0.3 \qquad\qquad \rho_{Tx} = \rho_{Rx} = 0.9$$
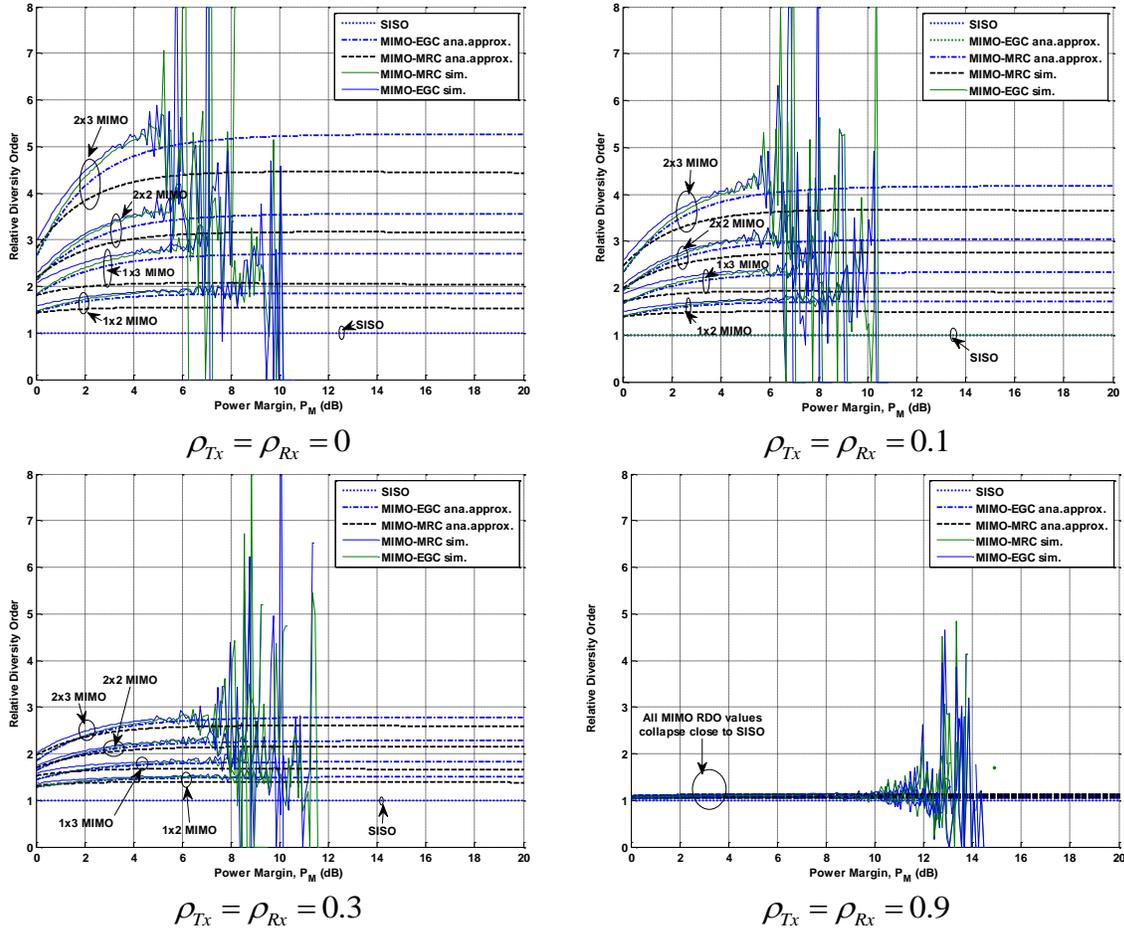
Fig. A.4 Relative Diversity Order versus Power Margin in dB for
$$\sigma_\chi^2 = 0.09, \rho_{Tx} = \rho_{Rx} = \{0, 0.1, 0.3, 0.9\}$$

Fig. A.4 shows the RDO values plotted for the above-mentioned correlation coefficient values for a log-normal variance of $\sigma_\chi^2 = 0.09$. The most outstanding finding is that the analytical approximation fails to hold for this large variance value. Both the EGC and MRC analytical approximations break down and diverge from the simulated values. So, caution must be exercised when employing analytical approximations for log-normal variables. The

approximation would only hold well for small values of *M* and *N*, and moderate values of σ_χ^2. This was obviously overlooked in the work by Safari and Uysal for $\sigma_\chi = 0.3$ [81].

However, the simulated results still show that the ARDO varies from a maximum of *MN* (for full correlation, $\rho_{Tx} = \rho_{Rx} = 0$) to a minimum of 1 (for full correlation, $\rho_{Tx} = \rho_{Rx} = 1$). Furthermore, there is no difference between the diversity order for MRC and EGC.

The conclusions drawn from this study are threefold: (1) for turbulence conditions exceeding a certain threshold, analytical approximation is not a good tool for closed-form analyses, (2) spatial correlation must be reduced to as practically low values as possible, by employing geometric configurations, so as to achieve the maximum benefit from multiple aperture systems, both for communications and imaging, and (3) MRC and EGC perform similarly in IM-DD systems. But we must also keep in mind that for a MIMO system, it is not practical to increase the number of transmitters and receivers beyond certain values, and these limits must be established and analytical approximations for these numbers must be matched up against simulation results to validate whether closed form expressions are adequately realistic.

REFERENCES

[1] DARPA Project SALTI on *http://space.hsv.usra.edu/CLRC/presentations/Tomlinson.ppt*

[2] M. Kavehrad, B. Hamzeh, "*Beaming Bandwidth via Laser Communications*", 5th Integrated Communications, Navigation and Surveillance Technologies Conference (ICNS), Fairfax, Virginia, May 2005.

[3] P. L. Fardley, D. R. Wisley, "*1Gbit/s optical free space link operating over 40m system and applications*", IEEE Proc. Optoelectron., Vol. 143, no. 6, pp. 330-333, Dec. 1996.

[4] H. Wu, M. Kavehrad, "*Availability Evaluation of Ground-to-Air Hybrid FSO/RF Links*", International Journal of Wireless Information Networks, Vol. 14, No. 1, pp. 33-45, Mar. 2007.

[5] X. Zhu, J. M. Kahn, "*Free-Space Optical Communication through Atmospheric Turbulence Channels*", IEEE Trans. On Communications, Vol. 50, No. 8, pp. 1293-1300, Aug. 2002.

[6] M. C. Roggemann, B. M. Welsh, "Imaging Through Turbulence", CRC Press, Boca Raton, Florida, 1996.

[7] G. Yun and M. Kavehrad, "*Spot diffusing and fly-eye receivers for indoor infrared wireless communications*", Proc. IEEE Int. Conf. Selected Topics in Wireless Communications, Vancouver, B. C., Canada, pp. 262-265, 1992.

[8] M. A. Albota, R. M. Heinrichs, D. G. Kocher, D. G. Fouche, B. E. Player, M. E. O'Brien, B. F. Aull, J. J.Zayhowski, J. Mooney, B. C. Willard, and R. R. Carlson, "*Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser*", Appl. Opt., Vol. 41, No. 36, pp 7671-7678, Dec. 2002.

[9] M.A. Albota, B.F. Aull, D. G. Fouche, R. M. Heinrichs, D. G. Kocher, R. M. Marino, J. G. Mooney, N. R. Newbury, M. E. O'Brien, B. E. Player, B. C. Willard, J. J. Zayhowski, "*Three-Dimensional Imaging Laser Radars with Geiger-Mode Avalanche Photodiode Arrays*", Lincoln Laboratory Journal, Vol. 13, No. 2, pp. 351–370, 2002.

[10] Brian F. Aull, Andrew H. Loomis, Douglas J. Young, Richard M. Heinrichs, Bradley J. Felton, Peter J. Daniels, and Deborah J. Landers, "*Geiger-Mode Avalanche Photodiodes for Three-Dimensional Imaging*", Lincoln Laboratory Journal, vol. 13, no. 2, pp. 335–350, 2002.

[11] S. Karp, R. Gagliardi, S. Moran, L. Stotts, *Optical Channels*, New. York: Plenum, 1988.

[12] A. Mannoni, C. Flesia, P. Bruscaglioni, and A. Ismaelli, "Multiple scattering from Chebyshev particles: Monte Carlo simulations for backscattering in lidar geometry," *Applied Optics*, Vol. 35, No. 36, pp. 7151-7164, Dec. 1996.

[13] F. G. Smith, *The Infrared & Electro-optical Systems Handbook: Atmospheric Propagating of Radiation*, Vol. 2, SPIE PRESS, 1993.

[14] Z. Hong, T. Demin, L. Lijuan, L. Yuelan, "Mie scattering characteristics of defects in cross linked polyethylene insulation of power cables," *Proceeding of 6th international conference on properties and applications of dielectric materials*, Vol. 1, pp. 529-533, Jun., 2000.

[15] B. Y. Hamzeh, *Multi-rate Wireless Optical Communications in Cloud Obscured Channels*, Ph.D. Thesis, Pennsylvania State University, 2005.

[16] D. Deirmendjian, "Scattering and Polarization Properties of Water Clouds and Hazes in the Visible and Infrared," *Appl. Opt*. Vol. 3, No. 2, pp 187-196, Feb. 1964.

[17] D. Deirmendjian, *Electromagnetic Scattering on Spherical Polydispersions*, New York: American Elsevier, 1969.

[18] T. Binzoni, T.S. Leung, A.H. Gandjbakhche, D. Rüfenacht, D. T. Delpy, "The use of the Henyey-Greenstein phase function in Monte Carlo simulations in biomedical optics," *Physics in Medicine and Biology*, Vol. 51, No. 17, pp. N313-N322, Sep. 2006.

[19] T. Binzoni, T.S. Leung, A.H. Gandjbakhche, D. Rüfenacht, D. T. Delpy, "Comment on 'The use of the Henyey-Grenstein phase function in Monte Carlo simulations in biomedical optics'," *Physics in Medicine and Biology*, Vol. 51, No. 22, pp. L39-L41(1), Nov. 2006.

[20] A. Sawetprawichkul, P. Hsu and K. Mitra, "Parallel Computing of Three-Dimensional Monte Carlo Simulation of Transient Radiative Transfer in Participating Media," *American Institute of Aeronautics and Astronautics, 8th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, St. Louis, Missouri, June 2002.

[21] L. Henyey and J. Greenstein, "Diffuse radiation in the galaxy," *Astrophysics Journal*, Vol. 93, pp. 70-83, 1941.

[22] R. C. Shirkey and D. H. Tofsted, "Electro-optical aerosol phase function database, PFNDAT2005," Army Research Laboratory, 2005.

[23] G. C. Mooradian, M. Geller, L. B. Stotts, D. H. Stephens, and R. A. Krautwald, "Blue-green pulsed propagation through fog," *Appl. Opt.* Vol. 18, pp. 429-441, Feb. 1979.

[24] J. R. Mourant, J. Boyer, A. H. Hielscher, I. J. Bigio, "Influence of the scattering phase function on light transport measurements in turbid media performed with small source–detector separations," *Opt. Lett*, Vol. 21, No. 7, pp. 546-548, April, 1996.

[25] Weisstein, Eric W. "Euler Angles," From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/EulerAngles.html.

[26] R. F. Lutomirski, A. P. Ciervo, and G. J. Hall, "Moments of multiple scattering," *Appl. Opt.* , Vol. 34, No. 30, 7125-7136, Oct. 1995.

[27] E. A. Bucher, "Computer simulation of light pulse propagation for communication through thick clouds," *Appl. Opt*, Vol. 12, No. 10, pp. 2391-2400. Oct. 1973.

[28] S. Chandrasekhar, *Radiative Transfer*, London: Oxford University Press, 1950.

[29] R. A. Dell-Imagine, *A Study of Multiple Scattering of. Optical Radiation With Applications to Laser Communications*, Ph.D. Dissertation in Engineering, University of California, Los Angles, 1965.

[30] H. M. Heggestad, "Optical Communications Through Multiple Scattering Media," MIT Tech. Report, No. 472, Nov. 22, 1968.

[31] H. M. Gupta, "Space-time response of a cloud communication channel to an optical signal*," Optical and Quantum Electronics*, Vol. 12, No. 6, pp 499-509, Nov. 1980.

[32] G. N. Plass and G. W. Kattawar, "Monte Carlo calculations of light scattering from clouds," *Appl. Opt.* , Vol. 7, No. 3, pp. 415–419, Mar. 1968.

[33] R. F. Lutomirski, A. P. Ciervo, and G. J. Hall, "Moments of multiple scattering," *Appl. Opt.* , Vol. 34, No. 30, 7125-7136, Oct. 1995.

[34] K. Riley, D. Ebert, M. Kraus, J. Tessendorf, and C. Hansen, "Efficient rendering of atmospheric phenomena," *Proceedings of Eurographics Symposium on Rendering 2004*, pp. 375-386, June 2004.

[35] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd Ed. New York: McGraw-Hill, 2002.

[36] Ehrhard Behrends, *Introduction to Markov chains with special emphasis on rapid mixing*, Berlin: Friedrick Vieweg & Son, 2001.

[37] S. Boyd, P Diaconis, and L. Xiao, "Fastest Mixing Markov Chain on a Graph," *SIAM Review*, Vol. 46, No. 4, pp. 667-689, December 2004.

[38] Binbin Wu, *Free-Space Optical Communications Through The Scattering Medium: Analysis of Signal Characteristics*, Ph.D. Thesis, Pennsylvania State University, December 2007.

[39] L. C. Andrews, R. L. Phillips, and C. Y. Hopen, *Laser Beam Scintillation with Applications*, SPIE Press, 2001, 1-66.

[40] R. Lane, A. Glindemann, and J. Dainty, "Simulation of a Kolmogorov phase screen," *Waves in Random meida*, Vol. 2, pp. 209-224, 1992.

[41] E. M. Johansson and D. T. Gavel, "Simulation of stellar speckle imaging," in *Amplitude and Intensity Spatial Interferometry II*, J. B. Breckinridge, ed., Proc. SPIE, Vol. 2200, pp 372–383, Mar. 1994.

[42] B. J. Herman and L. A. Strugala, ''Method for inclusion of low-frequency contributions in numerical representation of atmospheric turbulence,'' *Proc. SPIE.* Vol. 1221,pp. 183–192, Jan. 1990.

[43] R. L. Lucke, "Synthetic aperture ladar simulations with phase screens and Fourier propagation," *IEEE Aerospace Conference Proceedings* , pp. 1788-1798, Mar. 2004.

[44] P. Negrete-Regagnon, "Practical aspects of image recovery by means of the bispectrum," *J. Opt. Soc. Am.*, Vo. 13, No. 7, pp. 1557-1576, Jul. 1996.

[45] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* , Vol. 66, No. 3, pp. 207-211, Mar. 1976.

[46] M. Stoksik, R. G. Lane, and D. T. Nguyen, "Practical Synthesis of Accurate Fractal Images," *Graphical Models and Image Processing archive*, Volume 57, No. 3, pp 206-219, May, 1995.

[47] N. Roddier, "Atmospheric Wave-front Simulation Using Zernike Polynomials," *Optical Engineering*, Vol. 29, No. 10, pp 1174-1180, Oct. 1990.

[48] V. G. Orlov, V. V. Voitsekhovich, and S. Cuevas," Karhunen-Loeve Functions In Simulations of Atmospheric Distortions," *Revista Mexicana de Astronomia y Astrofisica,* Vol. 33, pp 187-195, 1997.

[49] D. L. Fried, "Probability of Getting a Lucky Short-Exposure Image Through Turbulence," *J. Opt. Soc. Am.*, Vol. 68, No. 12, pp 1651-1658, Dec. 1978.

[50] R. M. Marino, T. Stephens, R. E. Hatch, J. L. McLaughlin, J. G. Mooney, M. E. O'Brien, G. S. Rowe, J. S. Adams, L. Skelly, R. C. Knowlton, S. E. Forman, and W. R. Davis, "A compact 3D imaging laser radar system using Geiger-mode APD arrays: system and measurements," *Proc. SPIE*, Vol. 5086, pp 1-15, April 2003.

[51] Jaruwatanadilok, S., Ishimaru, A., and Kuga, Y., "Optical imaging through clouds and fog," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 8, pp 1834-43, Aug. 2003.

[52] Jivkova, S., and Kavehrad, M., "Multi-spot diffusing configuration for wireless infrared access: joint optimization of multi-beam transmitter and angle diversity receiver," *Proc. of SPIE International Symposium on Voice, Video, and Data Communications, Photonics East'99*, Vol. 3850, pp. 72-77, Sep.1999.

[53] Jivkova, S., and Kavehrad, M., "Indoor wireless infrared local access, multi-spot diffusing with computer generated holographic beam-splitter," *Proc. of the International Conference on Communications*, ICC'99, pp 604-608, June 1999.

[54] Jivkova, s., and Kavehrad, M., "Multi-spot diffusing configuration for wireless infrared access," IEEE Transactions on Communications, Vol. 48, No. 6, pp. 970-978, June 2000.

[55] IEC 825-1International Standard, Safety of laser products – Part 1: Equipment, classification, requirements and user's guide (1993).

[56] A. G. Basden, C. A. Haniff, and C. D. Mackay, "L3CCDs: fast photon counting for optical interferometry" Proc. Scientific Detectors Workshop, ASSL Library Series, 16-23 June 2002, Hawaii, Kluwer, 2002.

[57] J. W. Goodman, *Introduction to Fourier Optics*, 3rd Ed, Roberts & Company Publishers, 2005.

[58] B. C. McCallum, "Blind Deconvolution by Simulated Annealing", *Optics Communications*, Vol. 75, No. 2, pp. 101-105, Feb. 1990.

[59] D. Kundur, *Blind deconvolution of still images using recursive inverse filtering*, M.S. thesis, Univ. Toronto, Toronto, Ont., Canada, 1995.

[60] R. G. Lane and R. H. T. Bates, "Automatic multidimensional deconvolution," *J. Opt. Soc. Am.* , Vol. 4, pp.180–188, Jan. 1987.

[61] D. C. Ghiglia, L. A. Romero, and G. A. Mastin, "Systematic approach to two-dimensional blind deconvolution by zero-sheet separation," J. Opt. Soc. Am. A, vol. 10, pp. 1024-1036, May 1993.

[62] S. Gao, E. Kaltofen, J. May, Z. Yang, Lihong Zhi, "Approximate Factorization of Multivariate Polynomials via Differential Equations,", *International Conference on Symbolic and Algebraic Computation*, pp. 167-174, Santander, Spain, 2004.

[63] E. Kaltofen, J. May, Z. Yang, L. Zh, "Approximate Factorization of multivariate polynomials using singular value decomposition," *J. Symbolic Comput.*, Vol. 43, No. 5, pp.359-376, May 2008.

[64] D. Kundur and D. Hatzinakos, *Blind image deconvolution*, IEEE Signal Processing Magazine, vol. 13, pp. 43--64, May 1996.

[65] E. Simova, M. Tai, and M. Kavehrad, "Indoor wireless infrared link with a holographic multiple-spot diffuser", *Applications of Photonic Technology*, Vol. 2, Plenum Press, New York,1996, pp. 223-228.

[66] M. R. Feldman and C. C. Guest, "Iterative encoding of high-efficiency holograms for generation of spot arrays", *Optics Letters*, Vol. 14, No. 10, 1989, pp.479-481.

[67] M. P. Dames, R. J. Dowling, P. McKee, and D. Wood, "Efficient optical elements to generate intensity weighted spot arrays: design and fabrication", *Applied Optics*, Vol. 30, No. 19, 1991, pp. 2685-2691.

[68] P. Carnevali, L. Coletti, and S. Patarnello, "Image processing by simulated annealing", *IBM J. Res. Develop.*, Vol. 29, No.6, pp. 569-579, 1985.

[69] L. Solymar and D. J. Cooke, Volume holography and volume gratings, Academic Press, New York, NY, 1981.

[70] S. Jivkova and M. Kavehrad, "Multi-spot diffusing configuration for wireless infrared access: joint optimization of multi-beam transmitter and angle diversity receiver," *Proc. of SPIE International Symposium on Voice, Video, and Data Communications, Photonics East'99*, v. 3850, pp. 72-77, 1999.

[71] W. T. Welford and R. Winston, High collection non-imaging optics, Academic Press, San Diego, CA, 1989.

[72] J. B. Carruthers and J. M. Kahn, "Angle diversity for nondirected wireless infrared communication," *1998 IEEE Int. Conference on Communications*, 7-11 June 1998, Atlanta, Georgia, Conference Record, Vol. 3, pp.1665-1670.

[73] J. Kahn, R. You, P. Djahani, A. Weisbin, B. Teik, and A. Tang, "Imaging diversity receivers for high-speed infrared wireless communication," *IEEE Communications Magazine*, Vol. 36, No. 12, 1998, pp. 88-94.

[74] P. Tang, J. M. Kahn, and Keang-Po Ho, "Wireless infrared communication links using multi-beam transmitters and imaging receivers," *Proc. IEEE Int. Conference on Communications*, June 1996, Dallas, TX, pp.180-186.

[75] Keang-Po Ho and J. M. Kahn, "Compound parabolic concentrators for narrowband wireless infrared receivers," *Opt. Eng.*, Vol. 34, No. 5, 1995, pp. 1385-1395.

[76] S. Jivkova and M. Kavehrad, "Wireless infrared in-house communications: how to combat the multipath distortion," in *Optical Wireless Communication III,* E. J. Korevaar, ed., Proc. SPIE, **4214**, 162-170 (2000).

[77] V.Sainov, M.Mazakova and N.Koleva, "Characteristics of non-bleached reflection holograms," Comp. Rend. Bulg. Acad. Scien., **34**, 1241-44 (1981).

[78] V.Sainov, "Basic characteristics and application of reflection holograms," *Proc.of the 1-st Int. Symp. on Display Holography*, Lake Forest, Vol. 1, pp. 55-70 (1982).

[79] V.Sainov, S.Sainov and H. Bjelkhagen, "Color reflection holography," in *Practical Holography*, T. H. Jeong, J. E. Ludman, eds., Proc. SPIE, **615**, 88-92 (1986).

[80] K. Akhavan, M. Kavehrad and S. Jivkova, "Wireless infrared in-house communications: how to achieve very high bit rates," *WCNC'2000,* 23-28 September, Chicago, IL, Vol. 2, pp. 698 - 703.

[81] M. Safari, M. Uysal, "Diversity gain analysis of free-space optical communication systems", CCECE2008, 4-7 May 2008, Page(s): 1239-1244.

[82] S. M. Navidpour, M. Uysal, M. Kavehrad, "BER Performance of Free-Space Optical Transmission with Spatial Diversity", in *IEEE Trans. On Wireless Communications,* Vol. 6, No. 8, Aug. 2007.

[83] S. M. Navidpour and M. Uysal, M. Kavehrad, "*Capacity of MIMO lognormal Channels*", SPIE Photonic West, San Diego, CA, August 2006.

[84] S. Mohammad Navidpour, Murat Uysal and Mohsen Kavehrad, "*Performance Bounds for Correlated Turbulent Free-Space Optical Channels*", Proceedings of IEEE WCNC, Las Vegas, Nevada, April 2006.

[85] M. Safari, M. Uysal, "Relay-Assisted Free-Space Optical Communication", in *Proc. 41st ASILOMAR Conference on Signals, Systems and Computers 2007,* 4-7 Nov. 2007.

[86] Xiaoming Zhu, Kahn, J.M., "*Markov chain model in maximum-likelihood sequence detection for free-space optical communication through atmospheric turbulence channels*", IEEE Transactions on Communications, Volume 51, Issue 3, March 2003 Page(s):509 - 516.